

Tartu Ülikool

Matemaatika-informaatikateaduskond

Matemaatilise statistika instituut

Liina Tamme

Mitmese testimise probleem

Bakalaureusetöö (6 EAP)

Juhendaja: Marju Valge

Tartu 2014

Mitmese testimise probleem

Käesoleva töö eesmärk on anda ülevaade mitmese testimise probleemist. Analüüsides suure tunnuste arvuga andmestikke, puutume kokku paljude hüpoteesipaaride testimisega ehk mitmese testimisega. Mitmese testimise korral on vajalik meetodite kasutamine, et hoida I liiki vea tegemise tõenäosus kõikide testide korral alla soovitud olulise nivoo.

Töö esimeses peatükis antakse ülevaade mitmese testimise probleemist ning kirjeldatakse meetodeid, mis on keskendunud katseviisilisele veamäärale, tuues välja iga meetodi puudused ja head küljed. Samuti kirjeldatakse lühidalt seda, kuidas mõjutab testimist II liiki vea tegemine. Töö teises osas kirjeldatakse, kuidas kasutada esimeses osas tutvustatud meetodeid statistikapakettides R ja SAS.

Töö kolmandas osas rakendatakse kirjeldatud meetodeid andmestikule, mis koosneb Tartu Ülikooli 73 töötajagrupi keskmiste põhipalkade andmetest aastatel 2012 ja 2013. Analüüsi eesmärk on kindlaks teha, kas keskmine täiskoormusel põhipalk gruppides on aastaga muutunud. Mitmese testimise meetodeid kasutamata võetakse alternatiivne hüpotees palga muutumise kohta vastu grupis 70, mitmese testimise meetodeid kasutades ei võetud alternatiivset hüpoteesi vastu ühegi grupi puhul ja hoiti ära üks valepositiivne tulemus.

Märksõnad: statistiline analüüs, statistilised meetodid, testimine

Multiple Comparisons Problem

The purpose of this thesis was to give an overview of the multiple comparisons problem and to describe the multiple comparisons methods. It is typical that in a statistical analysis we come across with data that contains many variables and that we perform a large number of tests at the same time. In this case we need methods to hold the type I error below the desired level of significance.

In the first part of the thesis an overview of the multiple comparisons problem was given. Also methods that are focused on the family-wise error rate were described as well as the advantages and disadvantages of all the methods. A small overview was given of how type II error affects testing. In the second part of the thesis an overview of the functions to perform multiple comparison methods in statistical packages R and SAS was given.

In the last chapter of the thesis a multiple comparisons analysis was performed. Average full load salaries of the years 2012 and 2013 of Tartu University's 73 working groups were compared. The purpose of the analysis was to identify if the average full salary changed in a year. Without using the multiple comparison methods the salary difference of group 70 was declared significant. But when multiple comparison methods were used, all of them did not declare group 70 to be significant. Hence one false positive result was removed by using multiple comparison methods.

Keywords: statistical analysis, statistical methods, testing

Sisukord

1. SISSEJUHATUS	5
2. MITMENE TESTIMINE	6
2.1 ÜLEVAADE PROBLEEMIST	6
2.2 TÄHISTUSED	7
2.3 MEETODID	8
2.3.1 Bonferroni meetod	8
2.3.2 Dunni-Šidaki meetod.....	9
2.3.3 Bonferroni-Holmi meetod	10
2.3.4 Simesi-Hochbergi meetod	11
2.3.5 Hommeli meetod	12
2.4 NÄIDE MITMESE TESTIMISE MEETODITE KASUTAMISEST	13
2.5 TESTI VÕIMSUS	15
2.5 KOHANDATUD P-VÄÄRTUSED	17
2.6 NÄIDE KOHANDATUD P-VÄÄRTUSTE ARVUTAMISEKS	18
3. MITMESE TESTIMISE VÕIMALUSED STATISTIKAPAKETTIDES R JA SAS.. 21	
3.1 STATISTIKAPAKETI R VÕIMALUSED MITMESEKS TESTIMISEKS	21
3.1.1 Funktsiooni <i>p.adjust()</i> kirjeldus	21
3.1.2 Funktsiooni <i>pairwise.t.test()</i> kirjeldus	22
3.1.3 Näide kasutades R-i võimalusi mitmeseks testimiseks	23
3.2 STATISTIKAPAKETI SAS VÕIMALUSED MITMESEKS TESTIMISEKS	25
3.2.1 Näide kasutades SASi võimalusi mitmeseks testimiseks	28
4. TARTU ÜLIKOOLI TÖÖTAJATEGRUPPIDE KESKMISTE PÕHIPALKADE ANALÜÜS.....	30
4.1 ANDMESTIKU KIRJELDUS	30
4.2 TUNNUSTE KIRJELDUS.....	30
4.3 KESKMISTE PÕHIPALKADE VÕRDLUS TÖÖTAJAGRUPPIDES	31
4.4 KESKMISTE PÕHIPALGA MUUTUS	33
4.5 TARTU ÜLIKOOLI TÖÖTAJAGRUPPIDE KESKMISTE PÕHIPALGA STATISTILINE VÕRDLUS AASTATEL 2012 JA 2013	34
KASUTATUD KIRJANDUS	39
LISA 1. R-I KOOD NÄITE JAOKS.....	40
LISA 2. SAS-I KOOD NÄITE JAOKS	41
LISA 3. R-I KOOD ANALÜÜSI JAOKS	42

1. Sissejuhatus

Töö eesmärk on anda ülevaade mitmese testimise probleemist ning tutvuda meetoditega, mida mitmese testimise korral kasutada. On tavaline, et statistilises analüüsis puututakse kokku suure tunnuste arvuga andmestikega ning teostatakse korraga mitu erinevat testi. Mitmese testimise probleem seisneb selles, et korraga testitakse nii suurt hulka erinevaid hüpoteese, et mõni statistiliselt oluline seos leitakse isegi siis, kui seda tegelikult ei eksisteeri. Sel juhul on vajalik meetodite kasutamine, et hoida I liiki vea tegemise tõenäosus kõikide katsete korral alla soovitud olulise nivoo.

Töö esimeses osas tutvustatakse viit mitmese testimise meetodit: Bonferroni, Dunni-Šidaki, Bonferroni-Holmi, Simesi-Hochbergi ning Hommeli meetodit. Meetodite kirjeldamisel antakse ülevaade meetodi kasutamisevõimalustest ning meetodi puudustest.

Töö teises osas antakse ülevaade mitmese testimise meetodite kasutamisevõimalustest statistikapakettides R ja SAS. Kirjeldatakse R-i funktsioone `p.adjust()` ning `pairwise.t.test()`. SAS-s saab mitmest testimist teostada protseduuriga MULTTEST.

Töö kolmandas osas teostatakse analüüs Tartu Ülikooli 2012. ja 2013. aasta töötajagruppide keskmiste täiskoormusel põhipalkade võrdlemiseks. Tegemist on mitmese testimisega, kus võrreldakse 73 Tartu Ülikooli töötajagrupi 2012. aasta keskmist põhipalka 2013. aasta keskmise põhipalgaga.

Analüüs viidi läbi statistikapaketiga R. Bakalaureusetöö on kirjutatud tekstitöötlusprogrammiga Microsoft Word 2010.

Autor tänab käesoleva bakalaureusetöö juhendajat Marju Valget rohkete selgituste ja kasulike nõuannete eest.

2. Mitmene testimine

2.1 Ülevaade probleemist

Olgu meil n testitavat hüpoteeside paari ja olgu n_0 tõeste nullhüpoteeside arv. Tabelis 2.1 on näha nelja tundmatut suurust S , T , U ja V , millega puutume statistilisi teste tehes alati kokku. Iga hüpoteesi testimise jaoks määrame I liiki vea tegemise tõenäosuse, mis tabelis on tähistatud V -ga: tõenäosuse, et võtame vastu alternatiivse hüpoteesi, kui tegelikult kehtib nullhüpotees. Samuti saame vaadelda suurust R , mis on kõigi – nii tõeste kui väärade – vastu võetud alternatiivsete hüpoteeside arv. Tõenäosust jääda nullhüpoteesi juurde, kui tegelikult kehtib alternatiivne hüpotees, nimetatakse II liiki veaks (valenegatiivne tulemus) ja seda on tabelis tähistatud T -ga. Tõenäosust võtta vastu alternatiivne hüpotees, kui see ka tegelikult kehtib, on tabelis 2.1 tähistatud tähega S , tõenäosust jääda nullhüpoteesi kehtides selle juurde, tähega U .

Testi võimsus näitab, kui suur osa alternatiivsetest hüpoteesidest vastu võetakse, kui nad ka tegelikult tõesed on. Testi võimsus arvutatakse valemiga :

$$1 - T.$$

Testi võimsuse põhjal saame määrata, millised mitmese testimise meetodid on kasutamiseks paremad. Hüpoteese testides on meie eesmärgiks kontrollida V -d ja minimiseerida T . See tagab maksimaalse testi võimsuse hoides I liiki vea kontrolli all.

Tabel 2.1. Tundmatud suurused, millega puutume kokku statistiliste testide korral

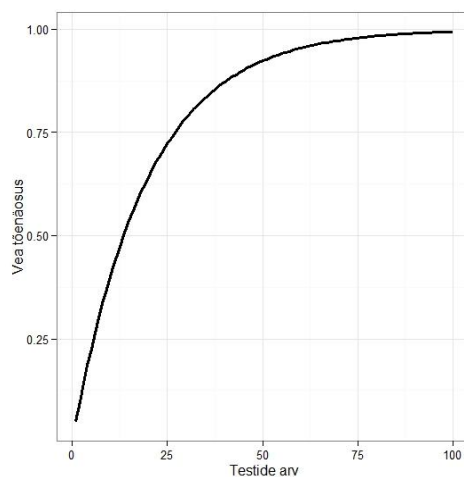
Tegelik olukord	Jäädud H_0 juurde	Vastuvõetud H_1	Kokku
Nullhüpotees (H_0)	U	V	n_0
Alternatiivne hüpotees (H_1)	T	S	$n - n_0$
Kokku	$n - R$	R	n

Esimest liiki viga, mis on tehtud vaid ühe hüpoteesipaari testimisel, nimetatakse võrdlusviisiliseks veamääraks (ingl *comparison-wise error rate*). Kogu katse peale tehtud esimest liiki viga nimetatakse katseviisiliseks veamääraks (ingl *family-wise error rate*), mida tähistatakse tavaliselt sümboliga π . Katseviisilist veamäära tähistatakse sageli ka lühendiga FWER.

Tõenäosus mitte teha I liiki viga n sõltumatu testi korral on $(1 - \alpha)^n$, kus α on testi olulisuse nivoo ja n testide arv. Olulisuse nivoo väljendab, kui suur statistiline viga on lubatud, kui väidame statistiliselt olulise seose olemasolu. Seega tõenäosus teha vähemalt üks I liiki viga on eespool toodud tõenäosuse vastandtõenäosus:

$$\pi = 1 - (1 - \alpha)^n.$$

Jooniselt 2.1 on näha, kui suur on vähemalt ühe I liiki vea tegemise tõenäosus eri testide arvu korral, kui testid on sõltumatud. Jooniselt näeme, et juhul, kui olulisuse nivoo on 0,05, siis ühe testi korral on vähemalt ühe I liiki vea tegemise tõenäosus 0,05, kuid kui teostatakse 100 testi, siis on tõenäosus teha vähemalt üks I liiki viga väga ühe lähedane. See tähendab, et peaaegu kindlalt võtame vastu alternatiivse hüpoteesi, mis tegelikult ei kehti.



Joonis 2.1. Vähemalt ühe I liiki vea tegemise tõenäosus juhul, kui olulisuse nivoo on 0,05

Mitmese testimise meetodite eesmärk on hoida katseviisiline veamäär π väiksemana soovitud olulisuse nivoost α .

2.2 Tähistused

Olgu meil n testi. Tähistame igale testile vastava teststatistiku T_i , kus $i = 1, \dots, n$. Juhusliku suuruse T_i realisatsiooniks olgu t_i , millele vastava p -väärtuse tähistame p_i . Juhul kui p -väärtused on järjestatud, tähistame i -nda p -väärtuse $p_{(i)}$ ja vastava teoreetilise teststatistiku $T_{(i)}$.

Testidele vastavaid hüpoteesipaare tähistame H_i . Kui testid on p -väärtuste alusel järjestatud, tähistame i -ndale p -väärtusele vastavat hüpoteesipaari kujul $H_{(i)}$.

2.3 Meetodid

2.3.1 Bonferroni meetod

Bonferroni meetod on oma nime saanud Itaalia matemaatiku Carlo Emilio Bonferroni järgi, sest meetodi konstrueerimisel on kasutatud Bonferroni võrratust. Tänapäevast kasutust peetakse Olive Jean Dunni poolt looduks (O. J. Dunn, 1959; O. J. Dunn, 1961).

Bonferroni meetod on arvatavasti vanim mitmese testimise meetod, mis on ka siimaani laialdaselt kasutusel. Bonferroni meetodi aluseks on Bonferroni võrratus, mille sisuks on, et tõenäosus teha I liiki viga kõikide testide korral, on väiksem või võrdne kui I liiki vea tegemise tõenäosuste summa iga testi korral eraldi.

Kui soovime, et kogu eksperimendi ulatuses I liiki vea tegemise tõenäosus oleks π , siis iga testi jaoks valime:

$$\alpha = \frac{\pi}{n}.$$

Nägemaks, et selline α valik tagab meile katseviisilise veamäära alla soovitud olulisuse nivoo, kasutame Boole'i võrratust:

$$P(\cup_i A_i) \leq \sum_i P(A_i), \text{ kus } A_i \text{ on loenduv arv sündmuseid.}$$

Olgu I_0 tõeste nullhüpoteeside hulk, millel on n_0 liiget. Katseviisiline veamäär on tõenäosus kummutada vähemalt üks nullhüpotees hulgast I_0 . Seega

$$P\left\{\cup_{i \in I_0} (p_i \leq \frac{\pi}{n})\right\} \leq \sum_{i \in I_0} \left\{P\left(p_i \leq \frac{\pi}{n}\right)\right\} \leq n_0 \frac{\pi}{n} \leq n \frac{\pi}{n} = \pi.$$

Bonferroni meetodi eelisteks on meetodi kasutamise lihtsus ning see, et meetodit saab kasutada nii sõltumatute kui ka sõltuvate testide korral.

Bonferroni meetodi puuduseks võib pidada tema liigset konservatiivsust. Testimisel võtame vastu alternatiivsed hüpoteesid vaid neil testidel, millele vastav p -väärtus on väiksem kui π/n .

kõigi ülejäänud testide korral jääme nullhüpoteesi juurde. Suure arvu testide korral tähendab see seda, et võrdleme teststatistikutele vastavaid p -väärtuseid väga väikese arvuga. See tähendab, et testi võimsus on väike.

Bonferroni meetod ei nõua, et testidele vastavaid p -väärtuseid võrreldaks võrdsete olulisuse nivooodega. See tähendab, et igale testile määratud olulisuse nivoo ei pea olema $\frac{\pi}{n}$, vajalik on, et nende summa kokku oleks π . See on otstarbekas juhul, kui mingi hüpoteesipaari vastuvõtmist eelistatakse teiste sisuliste hüpoteeside vastuvõtmisele (st analüüsi seisukohast on ühe alternatiivse hüpoteesi vastuvõtmine teistest olulisem). Näiteks juhul kui katseviisiline veamäär on 0,05, määrata eelistatud hüpoteesile võrdlusviisiliseks veamääraks 0,04 ja teistele $\frac{0,01}{n-1}$.

2.3.2 Dunni-Šidaki meetod

Dunni-Šidaki meetod on loodud 1967. aastal Zbyněk Šidák poolt (Šidák, 1967). See on lihtne ja konservatiivne meetod katseviisilise veamäära kontrollimiseks eeldades, et kõik testid on sõltumatud. Kasutades eespool tuletatud valemit $\pi = 1 - (1 - \alpha)^n$ (vt lk 7), avaldame α :

$$(1 - \alpha)^n = 1 - \pi$$

$$1 - \alpha = (1 - \pi)^{1/n}$$

$$\alpha = 1 - (1 - \pi)^{1/n}$$

See ongi Dunni-Šidaki meetodi korral tingimus α jaoks.

Dunni-Šidaki meetodi eeliseks on tema suurem võimsus võrreldes Bonferroni meetodiga. Kuna $\frac{\pi}{n} \leq 1 - (1 - \pi)^{1/n}$, kui $n > 1$, siis Dunn-Šidaki meetod annab meile tugevama tingimuse alternatiivse hüpoteesi vastuvõtmiseks. Samas on Dunn-Šidaki meetodi kasutamiseks vajalik testide sõltumatus, mida Bonferroni meetodi korral ei ole vaja eeldada.

Bonferroni ja Dunni-Šidaki meetodeid tuleks kasutada ainult juhtudel, kui testide arv on üsna väike, vastasel juhul toovad need endaga kaasa testi väga väikese võimsuse.

2.3.3 Bonferroni-Holmi meetod

Bonferroni-Holmi meetod on teise poole oma nimest saanud Sture Holmi järgi, kelle poolt on see meetod 1979. aastal loodud (Holm, 1979). Esimene pool nimest tuleneb sellest, et meetodi konstrueerimisel on kasutatud Bonferroni võrratust. Järgnev meetodi kirjeldus ning järeldused meetodi kohta põhinevad S. Holmi artiklil (Holm, 1979).

Bonferroni-Holmi meetodi korral järjestatakse testide p -väärtused kasvavas järjekorras:

$$p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(n)}.$$

Järgnevalt kasutatakse meetodit:

- Kui $p_{(1)} > \frac{\pi}{n}$, siis jäädakse nullhüpoteesi juurde iga testi korral.
- Kui $p_{(1)} \leq \frac{\pi}{n}$, siis võetakse vastu $H_{(1)}$ alternatiivne hüpotees ja vaadeldakse edasi $H_{(2)}$.
- Kui $p_{(2)} > \frac{\pi}{(n-1)}$, siis jäädakse nullhüpoteesi juurde iga testi $H_{(i)}$ korral, kus $i \geq 2$.
- Kui $p_{(2)} \leq \frac{\pi}{(n-1)}$, siis võetakse vastu $H_{(2)}$ alternatiivne hüpotees ja vaadeldakse edasi $H_{(3)}$.
- Protseduuri jätkatakse, kuni leidub j nii, et $p_{(j)} > \frac{\pi}{(n-j+1)}$, või kuni selgub, et kõik alternatiivsed hüpoteesid saab vastu võtta.

Tavalise Bonferroni meetodi korral võrreldakse hüpoteesipaaridele vastavaid p -väärtuseid suurusega π/n , kuid Bonferroni-Holmi meetodi korral suurustega $\pi/n, \pi/(n-1), \dots, \pi/1$.

See tähendab, et tõenäosus võtta vastu tõeseid alternatiivseid hüpoteese Bonferroni meetodi korral on väiksem või võrdne kui Bonferroni-Holmi meetodi korral. Bonferroni meetodi kasutamise saab alati asendada Bonferroni-Holmi meetodiga ilma, et väheneks tõenäosus võtta vastu tõeseid alternatiivseid hüpoteese.

Testi võimsuse suurenemine kasutades Bonferroni-Holmi meetodit Bonferroni meetodi asemel oleneb suuresti hüpoteesipaaridest. See on väike, kui kõik alternatiivsed hüpoteesid on „peaaegu õiged“ (st vastavad p -väärtused 0,05-lähedased) ning see on märgatav, kui mingi osa nullhüpoteesidest on „täiesti valed“ (st vastavad p -väärtused on 0-lähedased).

Bonferroni-Holmi meetodi suureks eeliseks (nagu ka Bonferroni meetodi korral) on tema paindlikkus. Puuduvad piirangud testide tüübile, mille korral meetodit võib kasutada, ainus tingimus on see, et peab olema võimalik arvutada vajalikke p -väärtuseid iga erineva testi jaoks.

Bonferroni-Holmi testi saab kasutada kõikides olukordades, kus Bonferroni meetoditki. Kuna Bonferroni-Holmi meetodi korral on arvutamismeetod vaid natuke keerulisem, aga testi võimsus arvestatavalt suurem, siis tuleks Bonferroni meetodi asemel kasutada Bonferroni-Holmi meetodit.

2.3.4 Simesi-Hochbergi meetod

Simesi-Hochbergi meetod on esimese poole oma nimest saanud R. J. Simesi järgi, kuna see põhineb meetodil, mida Simes 1986. aastal kirjeldas (Simes, 1986). Teine pool nimest tuleneb Yossi Hochbergi nimest, kelle poolt on meetod 1988. aastal loodud (Hochberg, 1988).

Simes pakkus 1986. aastal välja meetodi katseviisilise veamäära kontrollimiseks. Selle meetodi korral võetakse mistahes alternatiivne hüpotees vastu, kui $p_j \leq \frac{j\alpha}{n}$ vähemalt ühe $j = 1, \dots, n$ korral. Simes tõestas, et meetod kontrollib katseviisilist veamäära olulisuse nivool α , kui p -väärtused teststatistikutele on sõltumatud.

Simesi meetodi puuduseks on aga see, et puudub viis, kuidas teha järeldusi üksikute hüpoteeside jaoks. Selle probleemi lahendas Hochberg, kui pakkus välja modifitseeritud Bonferroni meetodi mitmese testimise jaoks, laiendades Simesi meetodit (Hochberg, 1988).

Simesi-Hochbergi meetodi korral järjestatakse testide p -väärtused kahanevas järjekorras:

$$p_{(n)} \geq p_{(n-1)} \geq \dots \geq p_{(1)}.$$

Järgnevalt kasutatakse meetodit:

- Kui $p_{(n)} \leq \pi$, siis iga testi korral võetakse vastu alternatiivne hüpotees.
- Kui $p_{(n)} > \pi$, siis $H_{(n)}$ ei saa kummutada ja vaadeldakse edasi $H_{(n-1)}$.
- Kui $p_{(n-1)} \leq \frac{\pi}{2}$, siis kõikide testide $H_{(i)}$, $i \leq n - 1$, korral võetakse vastu alternatiivne hüpotees.
- Kui $p_{(n-1)} > \frac{\pi}{2}$, siis $H_{(n-1)}$ ei saa kummutada ja vaadeldakse edasi $H_{(n-2)}$.

- Üldiselt, kui $p_{(n-i)} \leq \frac{\pi}{(n-i+1)}$, siis kõikide testide $H_{(i)}$, $i \leq n - i$, korral võetakse vastu alternatiivne hüpotees.

Simesi-Hochbergi meetodit tuleks eelistada juhtudel, kui testid on sõltumatud, sest siis on see Bonferroni-Holmi meetodist võimsam. Juhul kui testide sõltumatuses ei saa kindel olla, tuleks kasutada Bonferroni-Holmi meetodit.

Simesi-Hochbergi meetod lahendab olukorra, kus on näiteks kümme p -väärtust, mis kõik on võrdsed 0,04, hoides katseviisilist veamäära kontrolli all. Sellisel juhul jääksime Bonferroni ja Bonferroni-Holmi meetodite korral kõikide testide puhul nullhüpoteesi juurde, kuid Simesi-Hochbergi meetod kummutab kõik nullhüpoteesid.

2.3.5 Hommeli meetod

Hommeli meetod on oma nime saanud G. Hommeli järgi, kelle poolt on meetod 1988. aastal loodud (Hommel, 1988). Järgnev arutelu ning meetodi kirjeldus põhineb G. Hommeli artiklil (Hommel, 1988).

Teoreem. (Hommel, 1986) Olgu $n \geq 1$ testi jaoks antud n individuaalset hüpoteesi H_1, \dots, H_n ja defineerime $H_I = \bigcap_{i=1}^n H_i$ kõikide $I \in K$ jaoks, kus K on hulk, mis koosneb kõikidest mittetühjadest $\{1, \dots, n\}$ alamhulkadest. Oletame, et iga $I \in K$ jaoks eksisteerib test olulisuse nivool α teststatistikuga T_I . Võtame vastu hüpoteesi H_I , kui see on võetud vastu teststatistiku T_I poolt ja kui kõik H_J on vastu võetud T_J poolt, kus $J \supseteq I$, $J \in K$. Siis selline mitmese testimise meetod hoiab katseviisilise veamäära alla soovitud olulisuse nivoo α .

Kui lähtuda sellest printsiibist ja kasutada üksikute testide jaoks Simesi meetodit, saadakse järgnev mitmese testimise meetod, mida nimetatakse Hommeli meetodiks. Hommeli meetodi korral võetakse alternatiivne hüpotees vastu kõikide testide korral, mille puhul:

$$p_{(i)} \leq \frac{\pi}{k},$$

kus k on suurim täisarv, mille korral $p_{(n-k+j)} > \pi \frac{j}{k}$, ($j = 1, \dots, k$). Kui selline k ei eksisteeri, siis võetakse alternatiivne hüpotees vastu kõikide testide korral.

Meetod hoiab katseviisilise veamäära alla soovitud olulisuse nivoo juhul, kui testid on sõltumatud.

Nii Simesi-Hochbergi kui ka Hommeli meetod põhinevad Simesi meetodil, kuid Hommeli meetodi korral on testi võimsus suurem. Ühtlasi on Hommeli meetodi võimsus suurem või võrdne võrreldes kõikide eespool nimetatud meetoditega.

Võib öelda, et juhul, kui tegemist on sõltumatute testidega, tuleks kasutada Hommeli meetodit, kuna võrreldes teiste meetoditega on Hommeli meetodi võimsus kõige suurem. Sõltuvate testide korral on kõige suurema võimsusega Bonferroni-Holmi meetod.

2.4 Näide mitmese testimise meetodite kasutamisest

Olgu meil hüpoteesid H_1, \dots, H_{10} ning olgu p_1, \dots, p_{10} hüpoteesidele vastavate teststatistikute p -väärtused. Näites rakendame p -väärtustele kõiki eelnevalt kirjeldatud mitmese testimise meetodeid. Tabelis 2.2 on 10 p -väärtust esitatud kasvavas järjekorras ning lisame tabelisse ka Bonferroni-Holmi ning Simesi-Hochbergi meetodite rakendamise jaoks vajaliku suuruse $\frac{\pi}{n-i+1}$. Selle näite korral tahame, et kogu analüüsi ulatuses oleks I liiki vea tegemise tõenäosus 0,05. Seega $\pi = 0,05$.

Tabel 2.2 Näites kasutatavad p -väärtused

i	1	2	3	4	5	6	7	8	9	10
$p_{(i)}$	0,0010	0,0020	0,0045	0,0070	0,0085	0,0150	0,0500	0,0950	0,1550	0,3750
$\frac{\pi}{n-i+1}$	0,0050	0,0056	0,0063	0,0071	0,0083	0,0100	0,0125	0,0167	0,0250	0,0500

Järgnevalt viime läbi kõik eelnevalt kirjeldatud viis mitmese testimise meetodit, kasutades tabelis 2.2 toodud p -väärtusi.

Esimesena kasutame Bonferroni meetodit. Selleks peame kõiki p -väärtusi võrdlema suurusega $0,05/10 = 0,005$. Seega võrreldes etteantud p -väärtuseid suurusega 0,005 saame vastu võtta hüpoteesid $H_{(1)}, H_{(2)}$ ja $H_{(3)}$.

Dunni-Šidaki meetodi korral peame kõiki p -väärtusi võrdlema suurusega $1 - (1 - 0,05)^{0,1} = 0,0051$. Tulemuseks saame sama tulemuse, mis Bonferroni meetodi korral: saame vastu võtta hüpoteesid $H_{(1)}$, $H_{(2)}$ ja $H_{(3)}$.

Bonferroni-Holmi meetodi teostamiseks on p -väärtused juba kasvavalt järjestatud. Nüüd on vajalik võrrelda järjestatud p -väärtusi vastavalt olulisuse nivooga, mis on esitatud tabeli 2.2 kolmandas reas, $\frac{\pi}{n-i+1}$. Võrdlemist alustatakse kõige väiksemast p -väärtusest. Bonferroni-Holmi meetodi korral saame vastu võtta alternatiivsed hüpoteesid $H_{(1)}$, $H_{(2)}$, $H_{(3)}$ ja $H_{(4)}$, millele vastavad p -väärtused on $p_{(1)}$, $p_{(2)}$, $p_{(3)}$ ja $p_{(4)}$, sest $p_{(i)} \leq \frac{\pi}{n-i+1}$, kui $i \leq 4$. Protseduur lõpeb p -väärtuse $p_{(5)}$ võrdlemisel temale vastava olulisuse nivooga, sest $p_{(5)} = 0,0085 > 0,0083$.

Simesi-Hochbergi meetodi teostamiseks peame p -väärtuseid vaatlema kahanevas järjekorras, see tähendab, et vaatleme p -väärtuseid järjekorras $p_{(10)}, \dots, p_{(1)}$. Nüüd on vajalik võrrelda järjestatud p -väärtuseid vastavalt olulisuse nivooga, mis on esitatud tabeli 2.2 kolmandas reas, $\frac{\pi}{n-i+1}$. Simesi-Hochbergi meetodi korral saame vastu võtta alternatiivsed hüpoteesid $H_{(1)}$, $H_{(2)}$, $H_{(3)}$ ja $H_{(4)}$, millele vastavad p -väärtused on $p_{(1)}$, $p_{(2)}$, $p_{(3)}$ ja $p_{(4)}$, sest p -väärtuse $p_{(4)}$ korral on $p_{(4)} = 0,0070 \leq 0,0071$ ja seega võib vastu võtta kõik alternatiivsed hüpoteesid $H_{(i)}$, mille korral $i \leq 4$.

Hommeli meetodi korral võtame vastu kõik hüpoteesid, mille korral $p_{(i)} \leq \frac{\pi}{k}$, kus k on suurim täisarv, mille korral $p_{(n-k+j)} > \pi \frac{j}{k}$, ($j = 1, \dots, k$). Seega peame eelnevalt leidma suuruse k :

- Juhul kui $k = 1$ ja $j = 1$, siis $p_{(10)} = 0,3750 > 0,0500$.
- Juhul kui $k = 2$ ja $j = 1$, siis $p_{(9)} = 0,1550 > \frac{0,05}{2} = 0,0250$.
Juhul kui $k = 2$ ja $j = 2$, siis $p_{(10)} > 0,0500$.
- Juhul kui $k = 3$ ja $j = 1$, siis $p_{(8)} = 0,0950 > \frac{0,05}{3} = 0,0167$.
Juhul kui $k = 3$ ja $j = 2$, siis $p_{(9)} > \frac{0,05 \cdot 2}{3} = 0,0333$.
Juhul kui $k = 3$ ja $j = 3$, siis $p_{(10)} > 0,0500$.
- Juhul kui $k = 4$ ja $j = 1$, siis $p_{(7)} = 0,0500 > \frac{0,05}{4} = 0,0125$.
Juhul kui $k = 4$ ja $j = 2$, siis $p_{(8)} > \frac{0,05}{2} = 0,0250$.

Juhul kui $k = 4$ ja $j = 3$, siis $p_{(9)} > \frac{0,05 \cdot 3}{4} = 0,0375$.

Juhul kui $k = 4$ ja $j = 4$, siis $p_{(10)} > 0,0500$.

- Juhul kui $k = 5$ ja $j = 1$, siis $p_{(6)} = 0,0150 > \frac{0,05}{5} = 0,01$.

Juhul kui $k = 5$ ja $j = 2$, siis $p_{(7)} > \frac{0,05 \cdot 2}{5} = 0,02$.

Juhul kui $k = 5$ ja $j = 3$, siis $p_{(8)} > \frac{0,05 \cdot 3}{5} = 0,03$.

Juhul kui $k = 5$ ja $j = 4$, siis $p_{(9)} > \frac{0,05 \cdot 4}{5} = 0,04$.

Juhul kui $k = 5$ ja $j = 5$, siis $p_{(10)} > 0,05$.

- Juhul kui $k = 6$ ja $j = 1$, siis $p_{(5)} = 0,00850 > \frac{0,05}{6} = 0,0083$.

Juhul kui $k = 6$ ja $j = 2$, siis $p_{(6)} < \frac{0,05 \cdot 2}{6} = 0,0166$.

Seega selle näite korral on $k = 5$ ja võtame vastu kõik hüpoteesid, mille p -väärtus on väiksem või võrdne kui $\frac{0,05}{5} = 0,01$. Seega saame vastu võtta hüpoteesid $H_{(1)}, H_{(2)}, H_{(3)}, H_{(4)}$ ja $H_{(5)}$.

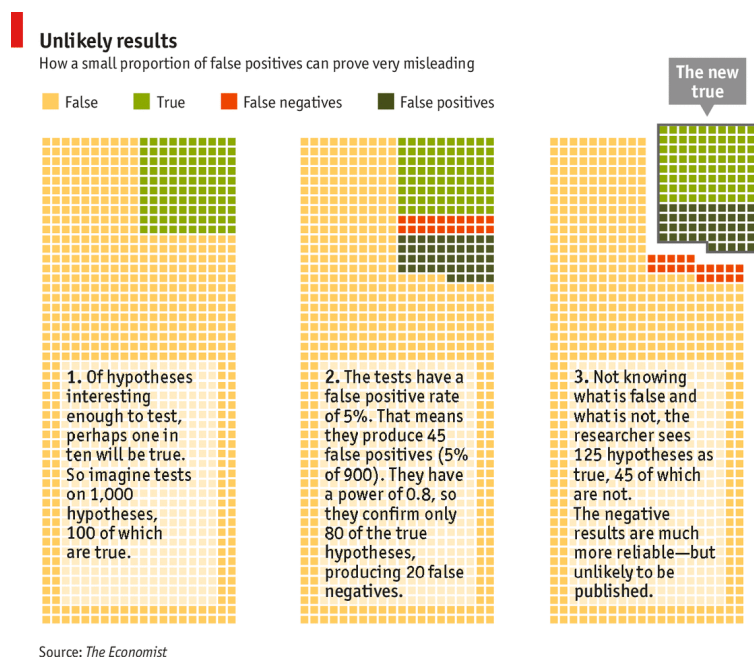
Seega kahe kõige konservatiivsema meetodi (Bonferroni meetodi ja Dunni-Šidaki meetodi) korral võeti vastu kõige vähem ehk 3 alternatiivset hüpoteesi. Bonferroni-Holmi ja Simesi-Hochbergi meetodite puhul võeti vastu 4 alternatiivset hüpoteesi ning kõige suurema võimsusega Hommeli meetodi korral võeti vastu enim ehk 5 alternatiivset hüpoteesi.

2.5 Testi võimsus

Testi võimsus näitab, kui suure osa tõestest alternatiivsetest hüpoteesidest test tuvastab. Seega juhul, kui testi võimsus on 0,8, siis 10 tõesest hüpoteesist tuvastab see test 8.

2013. aasta 19. oktoobri The Economisti väljaandes kirjutatakse artiklis „Trouble at the lab“ sellest, et teadlased keskenduvad oma töös liigselt I liiki vea hoidmisele alla soovitud olulise nivoo ning ei keskenduta piisavalt testi võimsusele.

Joonisel 2.2 on illustreeritud olukorda, kus testitakse 1 000 hüpoteesi, millest 100 juhul on alternatiivsed hüpoteesid tõesed. Kuna igas testis kasutatakse olulisuse nivood 0,05, siis 5% 900 testist (45 testi), mis ei ole statistiliselt olulised, loetakse oluliseks. Kui testi võimsus on 0,8, siis test loeb tõesteks 80 meie esialgsest 100 tõesest alternatiivsest hüpoteesist. Meile jääb alles 125 tõseks loetud alternatiivset hüpoteesi, millest 45 ehk 36% on tegelikult valed.



Joonis 2.2. Ülevaade testi võimsuse olulisusest testimisel¹

Artiklis mainitakse ka, et neuroteaduste alastes uurimustes on tüüpiline testi võimsus 0,21 ning psühholoogia-alastes uurimustes 0,35. Seega eelneva näite alusel loetakse neuroteaduste uurimustes tõseks 66 testi, millest 45 on valed, ning psühholoogia-alastes uurimustes loetakse tõesteks 80 testi, millest 45 on valed. Seega üle poolte tõseks loetud tulemustest on tegelikult valed.

Lisaks on teaduslikes artiklites suurem tendents avaldada tulemusi, mis on leidnud statistiliselt olulisi seoseid, see tähendab tulemused, kus võetakse vastu sisuline ehk alternatiivne hüpotees. Negatiivseid tulemusi avaldatakse 10–30% artiklitest.

¹ Joonis on leitav : <http://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble>

Tegelikuses on aga just negatiivsed tulemused palju usaldusväärsemad. Kasutades eespool toodud näidet, kus testi võimsus on 0,8, näeme, et meil on 875 negatiivset tulemust, millest ainult 20 ehk 2,3% on valed.

2.5 Kohandatud p -väärtused

Kui vaadelda üht hüpoteesi n hüpoteesi hulgast, siis kohandatud p -väärtuseks (ingl *adjusted p-value*) nimetatakse vähimat katseviisilist viga, mille korral selle teatud hüpoteesi korral alternatiivne hüpotees vastu võetakse. Kohandatud p -väärtust saab võrrelda otse valitud olulisuse nivooga. Kui kohandatud p -väärtus on väiksem või võrdne valitud olulisuse nivooga, siis võetakse vastu alternatiivne hüpotees.

Kohandatud p -väärtuse mõiste abil saab lihtsamini tõlgendada mitmese testimise meetodi tulemusi. Teststatistikutele vastavaid kohandatud p -väärtusi saab võrrelda ühe ja sama olulisuse nivooga, mitte enam iga testi jaoks erineva suurusega.

Järgnevalt anname kohandatud p -väärtused viiele eespool kirjeldatud meetodile. Nelja esimese kohandatud p -väärtuste leidmise valemid on pärit SASi kasutusjuhendist², kus kirjeldatakse mitmese testimise meetodeid.

Bonferroni meetodi korral on kohandatud p -väärtus i -nda testi jaoks võrdne

$$\tilde{p}_i = np_i, i = 1, \dots, n.$$

Kui kohandatud p -väärtus on suurem kui 1, siis võrdsustatakse see suurusega 1.

Dunni-Šidaki meetodi korral on kohandatud p -väärtus i -nda testi jaoks võrdne

$$\tilde{p}_i = 1 - (1 - p_i)^n, i = 1, \dots, n.$$

Bonferroni-Holmi meetodi korral on kohandatud p -väärtused $\tilde{p}_{(1)}, \dots, \tilde{p}_{(n)}$ võrdsed

$$\tilde{p}_{(i)} = \begin{cases} np_{(1)}, & \text{kui } i = 1 \\ \max(\tilde{p}_{(i-1)}, (n - i + 1)p_{(i)}), & \text{kui } i = 2, \dots, n \end{cases}$$

² Kasutusjuhend on leitav:

http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_multtest_sect014.htm#statug_multtest_multtestdhommel

Simesi-Hochbergi meetodi korral on kohandatud p -väärtused $\tilde{p}_{(1)}, \dots, \tilde{p}_{(n)}$ võrdsed

$$\tilde{p}_{(i)} = \begin{cases} p_{(n)}, & \text{kui } i = n \\ \min(\tilde{p}_{(i+1)}, (n - i + 1)p_{(i)}), & \text{kui } i = n - 1, \dots, 1 \end{cases}$$

Hommeli kohandatud p -väärtuste kohta on kirjutanud S. P. Wright (Wright, 1992), et Hommeli meetodi jaoks kohandatud p -väärtuste leidmiseks on vajalik kasutada Simesi meetodit ning kinnise testi protseduuri. Simesi meetodi korral teame, et n testi korral võetakse vastu hüpotees $H_0 = \{H_{(1)}, \dots, H_{(n)}\}$, millele vastavad p -väärtused $\{p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(n)}\}$, juhul kui $p_j \leq \frac{j\pi}{n}$ vähemalt ühe $j = 1, \dots, n$ korral. Sellest järeldub, et võetakse vastu H_0 juhul kui mingi $j = 1, \dots, n$ korral $\frac{np_j}{j} \leq \pi$. Seega hüpoteesi otsuse tegemisel on piisav teada minimaalne Simesi p -väärtus. Kinnise testi protseduur nõuab, et iga hüpoteesi H_i jaoks leitakse iga hüpoteesi sisaldava osakogumi jaoks, mis sisaldab hüpoteesi H_i , Simesi p -väärtus. Hommeli kohandatud p -väärtus on võrdne suurimaga neist p -väärtustest. Lisaks märgime, et pole vajalik arvutada Simesi p -väärtust kõikide alamhulkade jaoks, vaid iga m suurusest alamhulga jaoks on vajalik arvutada Simesi p -väärtused alamhulkade jaoks, mis sisaldavad meid huvitava hüpoteesi p -väärtust ning kõiki $(m - 1)$ järjekorras suuruselt järgmist p -väärtust. Näiteks, kui meil on 4 hüpoteesi ning me leiame Hommeli kohandatud p -väärtust hüpoteesi H_1 jaoks, siis on vajalik leida Simesi p -väärtused järgmiste alamhulkade jaoks: $A = (1, 2, 3, 4)$, $B = (1, 3, 4)$, $C = (1, 4)$ ning $D = (1)$.

2.6 Näide kohandatud p -väärtuste arvutamiseks

Olgu meil hüpoteesid H_1, \dots, H_5 ning olgu p_1, \dots, p_5 hüpoteesidele vastavate teststatistikute p -väärtused. Näites leiame kõikidele viiele eelnevalt kirjeldatud meetodile vastavad kohandatud p -väärtused. Kohandatud p -väärtuste leidmiseks kasutame olulisuse nivood 0,05. Tabelis 2.3 on esitatud p -väärtused, mis on järjestatud kasvavalt.

Tabel 2.3 Kohandatud p -väärtuste arvutamiseks kasutatavad p -väärtused

i	1	2	3	4	5
$p_{(i)}$	0,0010	0,0015	0,0060	0,0500	0,2500

Bonferroni meetodi kohandatud p -väärtused on seega võrsed suurustega $n \cdot p_{(i)} = 5 p_{(i)}$ ehk 0,0050, 0,0075, 0,0300, 0,2500 ja 0,7500. Kuna olulise nivoo on 0,05, siis järelikult võetakse vastu hüpoteesid $H_{(1)}$, $H_{(2)}$ ja $H_{(3)}$, mille korral $n \cdot p_{(i)} < 0,05$.

Dunni-Šidaki meetodi kohandatud p -väärtused arvutame valemi $1 - (1 - p_{(i)})^5$ põhjal. Seega Dunni-Šidaki meetodi kohandatud p -väärtused on vastavalt 0,0050, 0,0075, 0,0296, 0,2262 ja 0,7627. Seega vastu võetakse hüpoteesid $H_{(1)}$, $H_{(2)}$ ja $H_{(3)}$, mille korral kohandatud p -väärtused on väiksemad kui 0,05.

Bonferroni-Holmi meetodi korral arvutame kohandatud p -väärtuse eraldi iga $i = 1, \dots, 5$ korral:

- Juhul kui $i = 1$, siis $\tilde{p}_{(1)} = 0,0050$.
- Juhul kui $i = 2$, siis $\tilde{p}_{(2)} = \max(0,0050, 4 \cdot 0,0015) = 0,006$.
- Juhul kui $i = 3$, siis $\tilde{p}_{(3)} = \max(0,0060, 3 \cdot 0,0060) = 0,018$.
- Juhul kui $i = 4$, siis $\tilde{p}_{(4)} = \max(0,0180, 2 \cdot 0,0500) = 0,100$.
- Juhul kui $i = 5$, siis $\tilde{p}_{(5)} = \max(0,0100, 1 \cdot 0,2500) = 0,250$.

Seega vastu võetakse hüpoteesid $H_{(1)}$, $H_{(2)}$ ja $H_{(3)}$, mille korral kohandatud p -väärtused on väiksemad kui 0,05.

Simesi-Hochbergi meetodi korral arvutame kohandatud p -väärtuse eraldi iga $i = 1, \dots, 5$ korral:

- Juhul kui $i = 1$, siis $\tilde{p}_{(1)} = \min(0,0015, 5 \cdot 0,0010) = 0,005$.
- Juhul kui $i = 2$, siis $\tilde{p}_{(2)} = \min(0,0060, 4 \cdot 0,0015) = 0,006$.
- Juhul kui $i = 3$, siis $\tilde{p}_{(3)} = \min(0,0500, 3 \cdot 0,0060) = 0,018$.
- Juhul kui $i = 4$, siis $\tilde{p}_{(4)} = \min(0,2500, 2 \cdot 0,0500) = 0,100$.
- Juhul kui $i = 5$, siis $\tilde{p}_{(5)} = 0,250$.

Seega vastu võetakse hüpoteesid $H_{(1)}$, $H_{(2)}$ ja $H_{(3)}$, mille korral kohandatud p -väärtused on väiksemad kui 0,05.

Hommeli meetodi korral kohandatud p -väärtuste leidmiseks on koostatud parema ülevaate saamiseks tabel 2.4, kus on esitatud kõikide vajalike alamhulkade jaoks arvutatud Simesi p -väärtused ja seejärel leitud iga alamhulga jaoks vähim neist.

Tabel 2.4. Simes'i p -väärtused Hommeli kohandatud p -väärtuste leidmiseks

Alamhulga elemendid	$\frac{mp_i}{i}$	Vähim Simesi p -väärtus
1,2,3,4,5	0,0050, 0,0037, 0,0100, 0,0625, 0,2500	0,0037
1,3,4,5	0,0040, 0,0120, 0,0667, 0,2500	0,0040
1,4,5	0,0030, 0,0750, 0,2500	0,0030
1,5	0,0020, 0,2500	0,0020
1	0,0010	0,0010
2,3,4,5	0,0060, 0,0120, 0,0667, 0,2500	0,0060
2,4,5	0,0045, 0,0750, 0,2500	0,0045
2,5	0,0030, 0,2500	0,0030
2	0,0015	0,0015
3,4,5	0,0180, 0,0750, 0,2500	0,0180
3,5	0,012, 0,2500	0,0120
3	0,0060	0,0060
4,5	0,1000, 0,2500	0,1000
4	0,0500	0,0500
5	0,2500	0,2500

Kuna Hommeli kohandatud p -väärtus $\tilde{p}_{(i)}$ on võrdne suurima Simesi p -väärtusega, mis on arvutatud kõikide alamhulkade jaoks, mis sisaldavad i -ndat p -väärtust, siis:

$$\tilde{p}_{(1)} = \max(0,0037, 0,0040, 0,0030, 0,0020, 0,0010) = 0,004 \quad ,$$

$$\tilde{p}_{(2)} = \max(0,0037, 0,0060, 0,0045, 0,0030, 0,0015) = 0,006 \quad ,$$

$$\tilde{p}_{(3)} = \max(0,0037, 0,0060, 0,0180, 0,0120, 0,0060) = 0,018,$$

$$\tilde{p}_{(4)} = \max(0,0037, 0,0060, 0,0180, 0,1000, 0,0500) = 0,100 \quad ,$$

$$\tilde{p}_{(5)} = \max(0,0037, 0,0060, 0,0180, 0,1000, 0,2500) = 0,250.$$

Leitud Hommeli kohandatud p -väärtuste põhjal võtame vastu hüpoteesid $H_{(1)}$, $H_{(2)}$ ja $H_{(3)}$, mille korral kohandatud p -väärtused on väiksemad kui 0,05.

3. Mitmese testimise võimalused statistikapakettides R ja SAS

3.1 Statistikapaketi R võimalused mitmeseks testimiseks

Järgnevalt kirjeldatakse kahte võimalust, kuidas rakendada eelnevalt kirjeldatud mitmese testimise meetodeid. Esimene neist on R-i funktsioon `p.adjust()`, mis nõuab, et p -väärtused oleks eelnevalt leitud, ning seega saab seda kasutada kõikide analüüsimeetodite korral. Funktsioon `pairwise.t.test()` on kasutatav ainult juhul, kui analüüsi eesmärk on võrrelda keskmisi kasutades t -testi. Funktsiooni puhul ei ole vajalik eelnev p -väärtuste leidmine. Funktsioonide parameetrite ning väljastatavate tulemuste kirjelduse puhul on tegemist käesoleva töö autori tõlkega funktsioonide juhenditest³ ning lisatud on ka selgitavaid kommentaare.

3.1.1 Funktsiooni `p.adjust()` kirjeldus

Mitmese testimise korral saab eelnevaid meetodeid rakendada käsuga `p.adjust()`, mis andes ette p -väärtused, tagastab kohandatud p -väärtused vastavalt valitud mitmese testimise meetodile.

Funktsiooni `p.adjust()` üldkuju on järgmine:

```
p.adjust(p, method = p.adjust.methods, n = length(p))
```

Üldkujus välja toodud parameetrite tähendused:

`p` – etteantav eelnevalt leitud p -väärtuseid sisaldav vektor.

`method` – mitmese testimise meetod, millele saab anda väärtused „bonferroni“, „holm“, „hochberg“, „hommel“, mis tagastavad vastavalt Bonferroni, Bonferroni-Holmi, Simesi-Hochbergi ja Hommeli meetodi kohandatud p -väärtused.

`n` – teostatud võrdlemiste arv, mis peab olema vähemalt sama suur, kui p -väärtuste vektori pikkus. Kui võrdlemiste arv on suurem, siis Bonferroni ja Bonferroni-Holmi meetodi puhul eeldatakse, et vaatlemata jäävad p -väärtused on suuremad kõigist teistest vaadeldud p -väärtustest. Teiste meetodite korral eeldatakse, et vaatlemata p -väärtused on võrdsed ühega. Juhul, kui seda argumenti ette ei anta, kasutab funktsioon p -väärtuste vektori pikkust.

³ Juhendid on leitavad: <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/p.adjust.html> ning <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/pairwise.t.test.html>

Kuna selle funktsiooniga ei saa leida Dunni-Šidaki meetodile vastavaid kohandatud p -väärtuseid, lisatakse töö autori poolt kirjutatud funktsiooni nende leidmiseks.

```
sidak = function(p) {  
  
  tulemus = 1-(1 - p)**(length(p))  
  
  return(tulemus)  
  
}
```

Funktsiooni korral on vajalik esitada eelnevalt leitud p -väärtuste vektor, millele soovitakse Dunni-Šidaki meetodit rakendada. Funktsioon väljastab Dunni-Šidaki meetodi kohandatud p -väärtused.

3.1.2 Funktsiooni pairwise.t.test() kirjeldus

Mitmese testimise korral saab eelnevaid meetodeid rakendada funktsiooniga pairwise.t.test(), mis teostab paaridevahelised võrdlemised gruppide keskmiste vahel ja kasutab mitmese testimise meetodeid.

Funktsiooni pairwise.t.test() üldkuju on järgmine:

```
pairwise.t.test(x, g, p.adjust.method = p.adjust.methods,  
               pool.sd = !paired, paired = FALSE,  
               alternative = c("two.sided", "less", "greater"),  
               ...)
```

Üldkujus välja toodud parameetrite tähendused:

x – mõõtmistulemuste vektor.

g – vektor, mis sisaldab gruppide nimetusi.

$p.adjust.method$ – mitmese testimise meetod ehk p -väärtuse kohandamise meetod, millele saab anda väärtusteks „bonferroni“, „holm“, „hochberg“, „hommel“, mis tagastavad vastavalt Bonferroni, Bonferroni-Holmi, Simesi-Hochbergi ja Hommeli meetodi kohandatud p -väärtused.

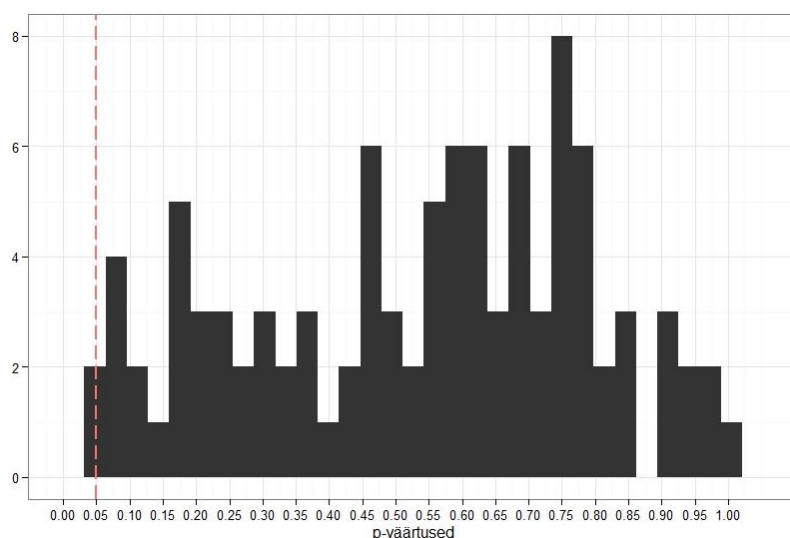
`pool.sd` – omandab väärtusi „TRUE“ ja „FALSE“, mis määravad, kas testimisel kasutatakse summaarset standardhälvet. Standardhälve summeerimine on mitme grupi standardhälve hindamise meetod, kus gruppide keskmised võivad olla erinevad, kuid võib eeldada, et standardhälbed on samad. See tähendab, et arvutatakse ühine standardhälve kõikide gruppide jaoks ja kasutatakse seda kõikide võrdlemiste jaoks (see on kasulik, kui mõned grupid on väga väikesed).

`paired` – omandab väärtusi „TRUE“ ja „FALSE“, mis määravad, kas testimisel kasutatakse paaris t-testi või mitte.

`alternative` – sõne, mis täpsustab alternatiivset hüpoteesi ja mille vaikeväärtuseks on „two.sided“ ehk kahepoolne hüpotees. Lisaks on ka võimalused „greater“ ja „less“, mis teostavad ühepoolseid hüpoteese. Kahepoolse hüpoteesi korral kontrollitakse, kas keskvärtus erineb, see tähendab, kas on väiksem või suurem mingist etteantud väärtusest. Ühepoolse testi võimaluse „greater“ korral kontrollitakse, kas keskvärtus on suurem mingist etteantud väärtusest ning võimaluse „less“ korral kontrollitakse, kas keskvärtus on väiksem mingist etteantud väärtusest.

3.1.3 Näide kasutades R-i võimalusi mitmeseks testimiseks

Käesolevas peatükis kirjeldatakse näite abil, kuidas kasutada R-i funktsiooni `p.adjust()`. Selleks genereeriti standardsest normaaljaotusest 200 valimit, igaühes 100 vaatlust. Seejärel jaotati valimid paarideks ning kasutati keskmiste võrdlemiseks t-testi (kokku 100 valimipaari). t-testi jaoks kasutati R-i funktsiooni `t.test()`, R-i kood selle teostamiseks on toodud lisas 1 (lk 40). t-testi teostamisel saadi 100 *p*-väärtust, mis on esitatud joonisel 3.1 (vt järgmiselt leheküljelt).



Joonis 3.1. p -väärtused enne, kui on kasutatud mitmese testimise meetodeid

Joonisele on lisatud punane vertikaalne joon, mis tähistab meid huvitavat väärtust 0,05. Saadud p -väärtuste põhjal peaksime ühe testi korral vastu võtma alternatiivse hüpoteesi ehk võtma vastu otsuse, et valimipaari keskmised ei ole võrdsed. Selle testi p -väärtus on 0,038. Ülejäänud testide korral on p -väärtused suuremad kui 0,05. Seega jääme nende testide korral nullhüpoteesi juurde, mis tähendab, et me ei saa tõestada, et valimipaaride keskmised erineksid. Saadud p -väärtused on arvatud kasutamata mitmese testimise meetodeid.

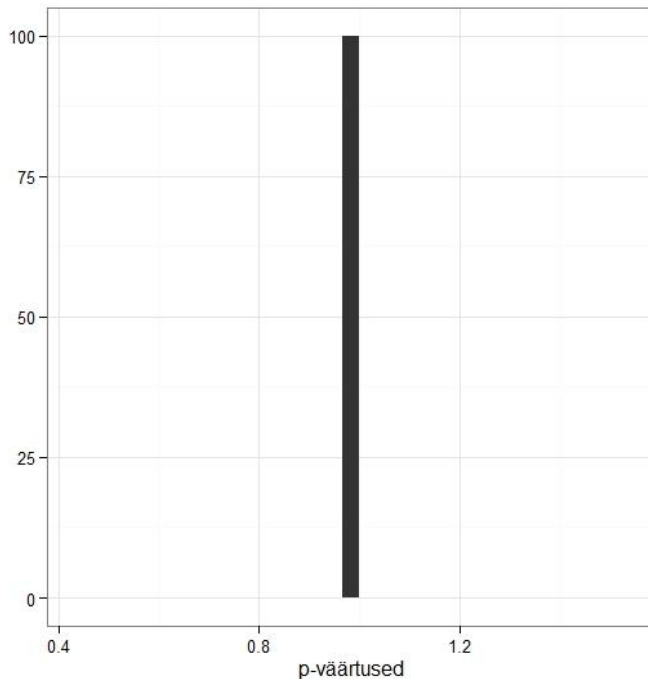
Mitmese testimise jaoks kasutati eespool kirjeldatud funktsiooni `p.adjust()`. Näide viidi läbi, kasutades Hommeli meetodit. R-i kood näite teostamiseks näeb välja järgnev:

```
p.adjust(pvals, method="hommel")
```

kus vektoris `pvals`, on eelnevalt välja arvutatud p -väärtused. Juhul, kui mitmeseks testimiseks tahetakse kasutada teisi meetodeid, on vajalik muuta muutuja „method“ väärtust. Näiteks, kui kasutatakse Bonferroni-Holmi meetodit, siis näeks R-i kood välja selline:

```
p.adjust(pvals, method="holm")
```

Hommeli meetodi kohandatud p -väärtused on esitatud joonisel 3.2 (vt järgmiselt leheküljelt).



Joonis 3.2. p -väärtused, kui on kasutatud Hommeli meetodit

Jooniselt näeme, et Hommeli meetodi kohandatud p -väärtuste korral jääme kõikide valimipaaride keskmiste võrdlemise korral nullhüpoteesi juurde.

3.2 Statistikapaketi SAS võimalused mitmeseks testimiseks

Mitmese testimise korral saab kirjeldatud meetodeid SAS-s rakendada protseduuriga MULTTEST. Protseduur väljastab sarnaselt R-i funktsioonidele samuti kohandatud p -väärtused. Protseduuri pole kohustuslik, kuid on võimalik kasutada juhul, kui p -väärtused on eelnevalt leitud. Funktsioonide parameetrite ning väljastatavate tulemuste kirjelduste puhul on tegemist autori tõlkega SASi kasutusjuhendist⁴ ning on lisatud ka selgitavaid kommentaare.

Protseduuri MULTTEST üldkuju on järgmine:

PROC MULTTEST <options> ;

BY variables ;

CLASS variable ;

CONTRAST 'label' values ;

⁴ Kasutusjuhend on leitav:

http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_multtest_sect005.htm

FREQ variable ;

STRATA variable ;

TEST name (variables </ options>) ;

Protseduurilause alustab protseduuri MULTTEST ja täpsustab kasutatava p -väärtuse kohandamise meetodi. Järgnevas tabelis 3.1 on täpsustatud võimalikud käsud soovitud meetodite kasutamiseks, mis saab ette anda protseduurilauses.

Tabel 3.1 Võimalikud käsud protseduurilauses

Võimalus	Kirjeldus
BONFERRONI	Arvutab Bonferroni kohandatud p -väärtused.
SIDAK	Arvutab Dunni-Šidaki kohandatud p -väärtused.
HOLM	Arvutab Bonferroni-Holmi kohandatud p -väärtused.
HOCHBERG	Arvutab Simesi-Hochbergi kohandatud p -väärtused
HOMMEL	Arvutab Hommeli kohandatud p -väärtused.
Sisestatavad ja väljastatavad andmestikud	
DATA=	Nimetab SASi andmestiku, mida kasutatakse protseduuriga MULTTEST.
INPVALUES=	Nimetab SASi andmestiku, mis koosneb p -väärtustest, mis on saadud mitmese testimise meetodeid kasutamata.
OUT=	Nimetab väljastatava SASi andmestiku, mis sisaldab kõiki muutujate nimesid, kontrastide nimesid, vahepealseid arvutusi ja kõiki p -väärtuseid ning kohandatud p -väärtuseid.
OUTSAMP=	Nimetab väljastatava andmestiku, mis sisaldab informatsiooni juhuvalikuga saadud andmetest juhul, kui on teostatud juhuvalik
Väljundi (OUTPUT) võimalused	
NOPRINT	Ei väljasta ühtegi tabelit.
NOTABLES	Ei väljasta diskreetsete muutujate tabelit ja pidevate muutujate tabelit.
NOZEROS	Ei väljasta tabeleid, kus esinevad nullid kõikidel grupi tasemetel.
NOPVALUE	Ei väljasta p -väärtuste tabelit, mis koosneb tavalistest p -väärtustest ja kohandatud p -väärtustest.
PLOTS	Väljastatakse ODS graafik.
Arvutuslikud võimalused	
EPSILON=	Täpsustab, kui suur peab olema kahe p -väärtuse erinevus, et need oleksid tuvastatud kui mittevõrdsed. Vaikeväärtuseks SAS 9.1 versioonis ja vanemates on 10^{-8} , uuemates versioonides 10^{-12} .

Lause „BY“ võimaldab teostada eraldi analüüsid erinevates gruppides, mis on määratud BY muutujatega. Kui kasutada seda võimalust, siis on vajalik, et kõik kasutatavad muutujad oleks eelnevalt sorteeritud.

Lause „CLASS“ on vajalik, kui protseduurilauses pole täpsustatud sisestust „INPVALUES“. Lause täpsustab ühe muutuja, mille põhjal jaotatakse andmestik gruppidesse, ja selle põhjal teostatakse analüüs. Muutuja võib olla nii numbriline kui ka kvalitatiivne.

Lause „CONTRAST“ täpsustab, millised võrdlemised teostatakse gruppides, mis on ette antud lauses „CLASS“. Sõne „label“ annab kontrastidele pealkirja. Teostades Fisheri täpset testi, on kontrasti koefitsiendid -1, 0 ja 1. Omavahel võrreldakse gruppe, mille kontrasti koefitsiendid on 1 ja -1. Kui kontrasti koefitsient on väärtusega 0, siis seda võrdlemisel ei kasutata. Kui ei täpsustata lauset „CONTRAST“, siis teostatakse Fisheri täpse testi korral kõik omavahelised võrdlemised. Muude testide korral on kontrasti väärtused vaikesi 0,1,2

Lause „FREQ“ nimetab muutuja, mis annab ette tihedused igale vaatlusele, mis sisalduvad ette antud andmestikus. Täpsemalt, kui n on muutuja FREQ väärtus mingile teatud vaatlusele, siis seda vaatlust kasutatakse n korda.

Lause „STRATA“ nimetab muutuja, mida kasutatakse analüüsis kihistamise muutujana. Lisades pärast lauset „STRATA“ kaldkriipsu (/) saab kasutada võimalust „WEIGHT“. See täpsustab kihtide kaalumise tüübi, mida kasutatakse, kui arvutatakse Freeman-Tukey või t-testi statistikuid.

Lause „TEST“ on vajalik, kui pole täpsustatud sisestust „INPVALUES“. Lause nimetab kasutatava testi ning kasutatavad diskreetsed ja pidevad väärtused. Võimalikud väärtused testidele on esitatud järgnevas tabelis 3.2.

Tabel 3.2. Võimalikud testid, mida saab kasutada protseduuris MULTTEST

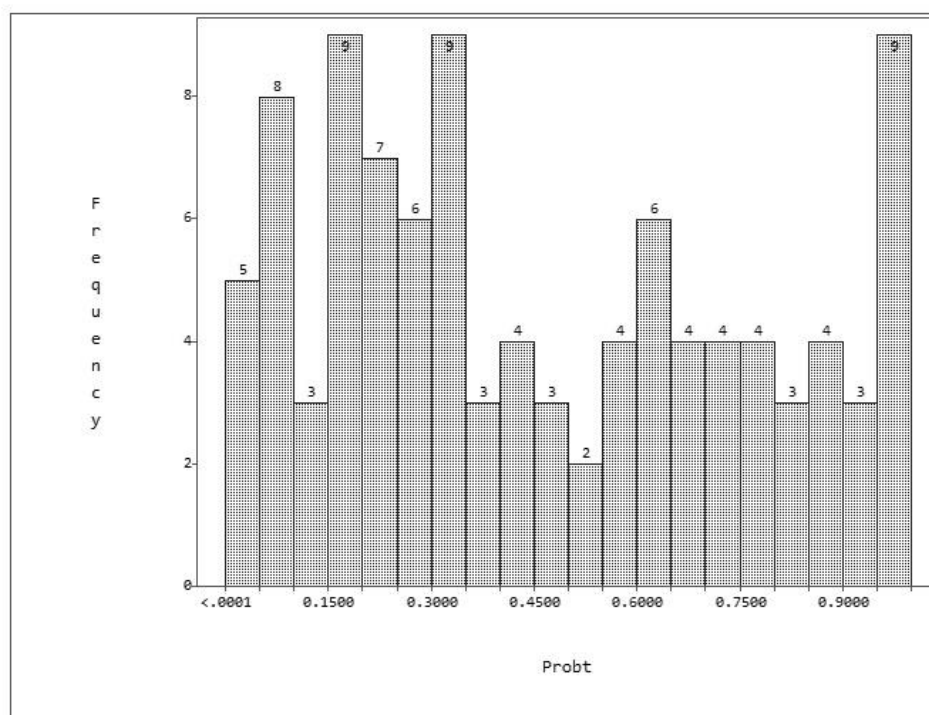
Testi nimetus	Kirjeldus
FISCHER	Teostab Fisheri täpse testi, mis võrdleb kaht ravigruppi. Testi muutujad peavad olema väärtusega „1“, kui on tegemist õnnestumisega, ning „0“ vastasel korral.
MEAN	Teostab t-testi keskmiste võrdlemiseks.

Testi lause juures saab täpsustamiseks kasutada võimalusi „LOWERTAILED“ või „UPPERTAILED“, mis muudavad kõik teostatavad testid ühepoolseteks. Võimaluse „LOWERTAILED“ korral kontrollitakse, kas parameeter on väiksem mingist etteantud väärtusest. Võimaluse „UPPERTAILED“ korral kontrollitakse, kas parameeter on suurem mingist etteantud väärtusest. Kui neid võimalusi ei kasutata on vaikumisi väärtus kasutada kahepoolset testi.

3.2.1 Näide kasutades SASi võimalusi mitmeseks testimiseks

SAS-is teostati näitena samasugune analüüs nagu R-i võimalusi kasutadeski. Seega genereeriti standardsest normaaljaotusest 200 valimit, igas valimis 100 objekti. Seejärel jaotati 200 valimit 100 grupiks, igas grupis 2 valimit. Igas grupis teostati t-test keskmiste võrdlemiseks. Näites kasutati p -väärtuste kohandamiseks Hommeli meetodit. SASi kood näite teostamiseks on esitatud lisas 2 (lk 41).

Joonisel 3.3 on esitatud t-testide teostamisel saadud p -väärtused.



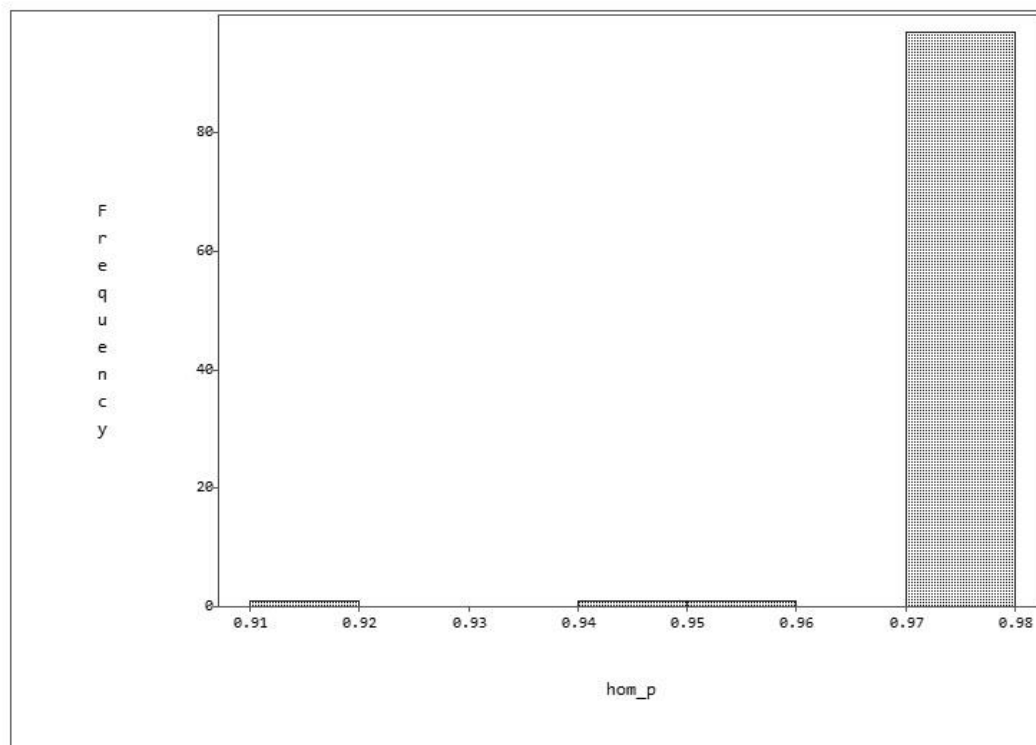
Joonis 3.3. Keskmiste võrdlemisel saadud p -väärtused kasutamata mitmese testimise meetodeid

Jooniselt näeme, et p -väärtus oli viies grupis väiksem kui 0,05 ehk nende gruppide korral peaks vastu võtma alternatiivse hüpoteesi, et keskmised on erinevad.

Järgmisena kasutame SAS-i protseduuri MULTTEST, et arvutada saadud p -väärtustele vastavad Hommeli kohandatud p -väärtused. Etteantud andmestik sisaldab 100 p -väärtust, mis on eelnevalt leitud. SAS-i kood selle jaoks näeb välja järgmine:

```
proc multtest INPVALUES=andmestik hommeli ods=results;  
  
run;
```

Tulemusena väljastatakse Hommeli kohandatud p -väärtused. Tulemused on esitatud joonisel 3.4.



Joonis 3.4. Keskmiste võrdlemisel saadud kohandatud p -väärtused kasutades Hommeli meetodit

Jooniselt näeme, et kasutades Hommeli meetodit, jääme iga valimite grupi keskmiste võrdlemisel nullhüpoteesi juurde ehk selle juurde, et keskmised on võrdsed.

4. Tartu Ülikooli töötajategruppide keskmiste põhipalkade analüüs

4.1 Andmestiku kirjeldus

Andmestik koosneb Tartu Ülikooli töötajategruppide 2012. ja 2013. aasta sissetuleku andmetest. Andmestikus on 4208 erineva töötaja andmed ning 4 tunnust. Mõne töötaja andmed esinesid andmestikus mõlemal aastal ning seega on andmestikus kokku 7751 rida. Töötajagruppe on kokku 73 ning igal töötajal grupis on oma unikaalne ID. Andmete konfidentsiaalsuse tõttu on andmestikus töötajagrupid märgitud numbriga ning pole teada, mis selle taga seisab. Lisaks ei avaldata töös töötajagruppi kuuluvate töötajate arvu. Aastal 2012 oli töötajaid vaatlusalustes gruppides kokku 3 897 ning 2013. aastal 3 854. Sissetuleku andmed on võetud aasta lõpu seisuga ning on taandatud täiskoormusele.

Andmestikus on järgnevad tunnused:

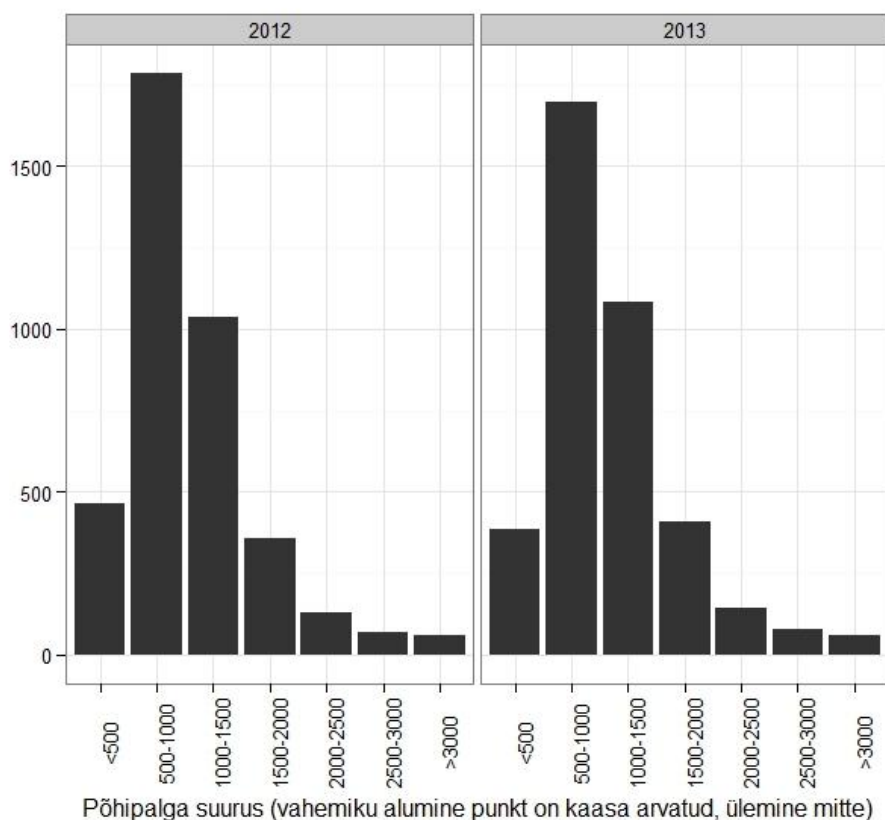
- **Aasta** – kaks väärtust (2012 ja 2013).
- **Töötajagrupi ID** – andmeid on kokku 73 erineva TÜ töötajagrupi kohta, mis on tähistatud numbrita 1, 2, 3, ..., 73.
- **Töötaja ID** – unikaalne töötaja ID.
- **Täiskoormusel põhipalk** – töötaja täiskoormusele taandatud põhipalk aasta lõpu (31. detsember) seisuga.

4.2 Tunnuste kirjeldus

Aastal 2012 oli kõigi vaatluse alla kuuluvate töötajate keskmine põhipalk 1 059,9 eurot ning aastal 2013 oli sama näitaja 1 119,3 eurot. Seega tõusis Tartu Ülikooli töötajate keskmine sissetulek aastaga 5,6%. Aastal 2012 oli minimaalne sissetulek 290 eurot ning aastal 2013 oli see näitaja 320 eurot. Need on ka vastavalt aastatel 2012 ja 2013 kehtinud miinimumpalgad. Aastal 2012 said miinimumpalka 37 töötajagruppidesse kuuluvat töötajat ning 2013. aastal 39 töötajat.

Joonisel 4.1 (vt järgmisel leheküljelt) on esitatud palkade jaotus aastatel 2012 ja 2013. Jooniselt võib näha, et kõige suurema osa moodustavad töötajad, kes saavad palka vahemikus 500–1000 eurot. Palkade jaotus pole üldjoontes aastaga muutunud, küll aga võib näha, et 2013. aastal on vähenenud töötajate hulk, kes saavad palka alla 500 euro või täpselt 500 eurot, ning töötajate arv, kes saavad palka vahemikus 500 – 1000 eurot (vastavalt 11,9%-lt 10,0%-le

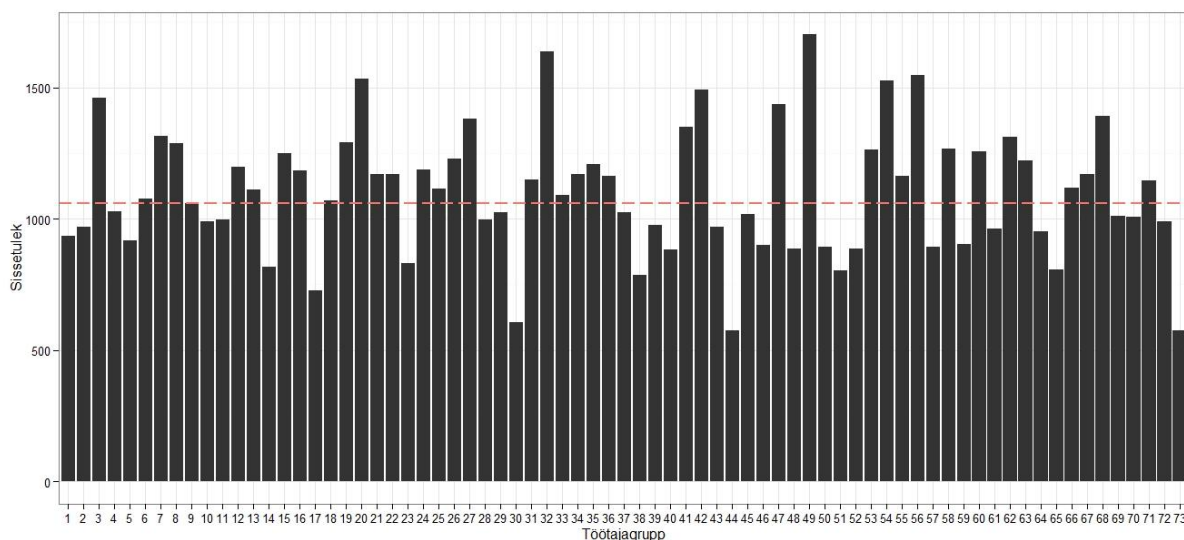
ning 45,8%-lt 44,0%-le). Samas on tõusnud töötajate hulk, kes saavad palka vahemikus 1000–3000 eurot. Töötajate osakaal, kes saavad palka üle 3000 euro on mõlemal aastal 1,5%.



Joonis 4.1. Palkade jaotus aastatel 2012 ja 2013

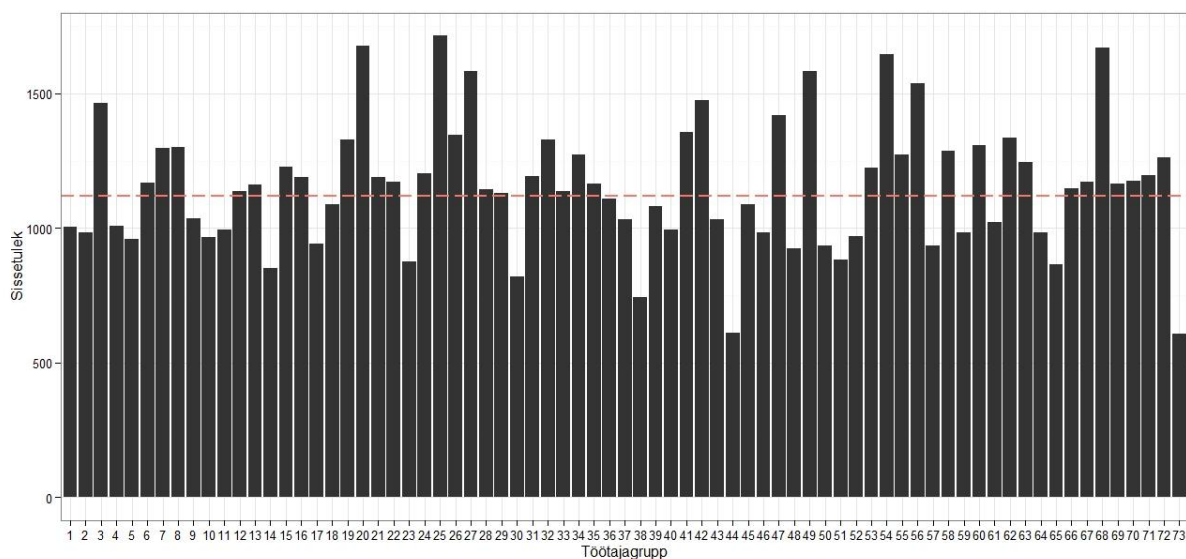
4.3 Keskmiste põhipalkade võrdlus töötajagruppides

Joonisel 4.2 (vt järgmiselt leheküljelt) on esitatud 2012. aasta keskmised sissetulekud kõikides vaadeldavates töötajagruppides. Joonisele on lisatud punase horisontaalse joonega 2012. aasta keskmine täiskoormusel põhipalk üle kõikide gruppide, mille väärtuseks on 1 059,9 eurot. Jooniselt võib näha, et kõrgeim keskmine põhipalk aastal 2012 kuulub töötajagrupile 49, kus keskmine põhipalk oli 1 701,9 eurot. Lisaks grupile 49 oli keskmine sissetulek üle 1 500 euro veel neljas grupis. Need grupid olid 54, 20, 56 ja 32, kus keskmised põhipalgad olid vastavalt 1 525,5 eurot, 1 535,0 eurot, 1 547,6 eurot ning 1 636,4 eurot. Kõige madalam keskmine põhipalk aastal 2012 oli grupis 44, mille keskmise palga näitaja oli 574,7 eurot. Võrreldes Tartu Ülikooli kõigi töötajagruppide keskmise palgaga 2012. aastal, oli keskmine põhipalk märkimisväärselt madal ka gruppides 73, 30 ja 17, kus vastavad näitajad olid 576,8 eurot 606,5 eurot ning 726,3 eurot.



Joonis 4.2. Töötajagruppide keskmine täiskoormusel põhipalk aastal 2012

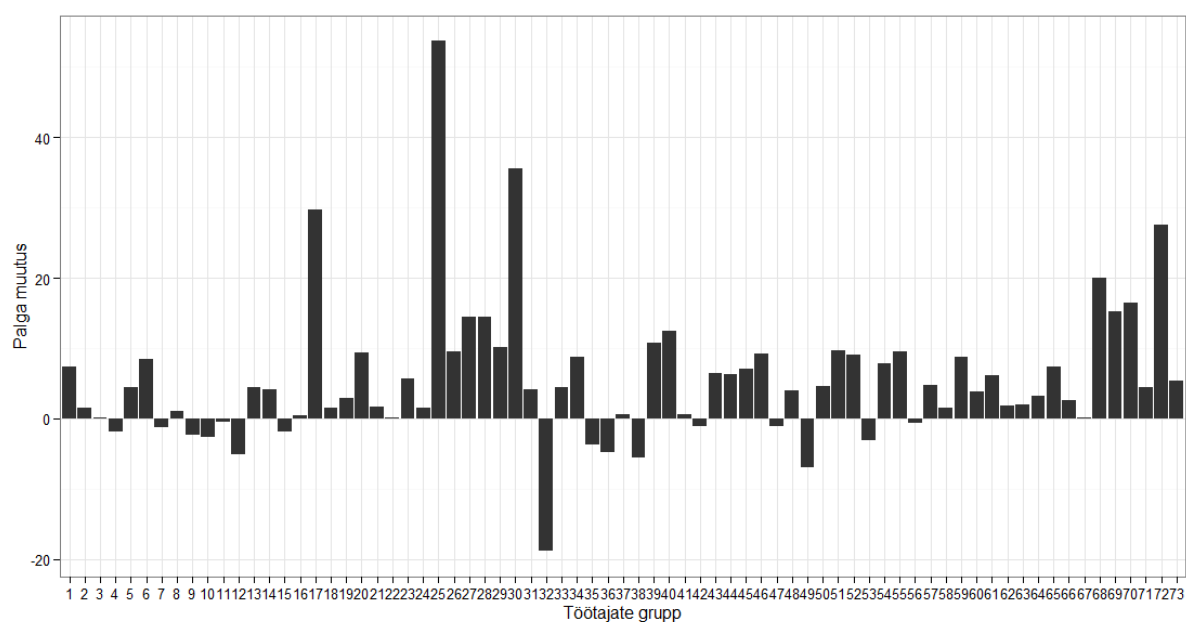
Joonisel 4.3 (vt järgmiselt leheküljelt) on esitatud 2013. aasta keskmised täiskoormusel põhipalgad kõikides töötajagruppides. Joonisele on lisatud punase horisontaalse joonega 2013. aasta keskmine täiskoormusel põhipalk üle kõikide gruppide, mille väärtuseks on 1 119,3 eurot. Jooniselt võib näha, et keskmine sissetulek oli aastal 2013 kõrgeim töötajagrupis 25, kus keskmine põhipalk oli 1 715,0 eurot. Lisaks grupile 25 oli keskmine sissetulek üle 1 500 euro veel kuues grupis. Need grupid olid 20, 27, 49, 54, 56 ja 68, kus keskmised põhipalgad olid vastavalt 1 679,1 eurot, 1 582,0 eurot, 1 582,8 eurot, 1 645,7 eurot, 1 537,0 eurot ning 1 671,0 eurot. Seega gruppide hulka, kes teenivad keskmiselt üle 1 500 euro, lisandusid aastaga grupid 25, 27, 54 ja 68. Samas ei saa keskmist palka üle 1 500 euro enam grupi 32 töötajad. Kõige madalam oli keskmine palk aastal 2013 grupis 73, mille vastav näitaja oli 608,0 eurot. Võrreldes Tartu Ülikooli kõigi töötajagruppide keskmise palgaga 2013. aastal, oli keskmine põhipalk märkimisväärselt madalam ka gruppides 44 ja 38, mille vastavad näitajad olid 742,5 eurot ning 611,0 eurot.



Joonis 4.3. Töötajagruppide keskmine täiskoormusel põhipalk aastal 2013

4.4 Keskmise põhipalga muutus

Joonisel 4.4 on esitatud iga töötajagrupi keskmiste põhipalkade muutus võrreldes aastaid 2012 ja 2013. Jooniselt näeme, et suuremas osas töötajagruppides (78%) keskmine palk suurenes. Samas oli ka töötajate grappe, kus keskmine põhipalk vähenes. Neid oli kokku 16 ehk 22% kõigist gruppidest.



Joonis 4.4 Töötajagruppide keskmise põhipalga muutus 2013. aastal võrreldes aastaga 2012

Kõige suurem oli keskmise palga tõus gruppis 25, kus keskmine palk tõusis 53,7%. Üle neljandiku tõusid palgad ka veel gruppides 72, 17 ja 30 (vastavalt 27,5%, 29,6% ja 35,5%). Kõige rohkem langes keskmine palk gruppis 32 (18,8%). Üle 5% langes palk ka gruppides 49, ja 38 (vastavalt 7% ja 5,6 %).

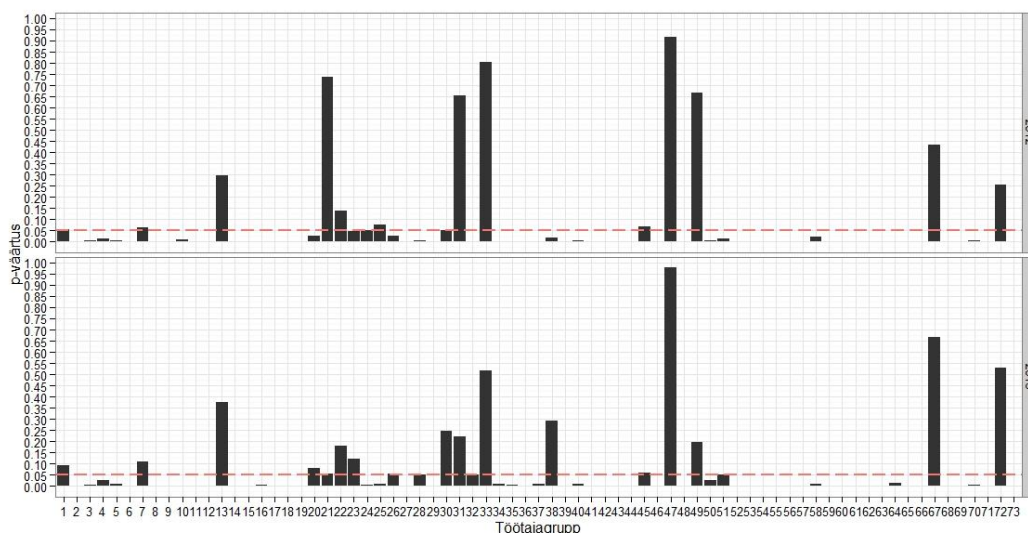
4.5 Tartu Ülikooli töötajagruppide keskmise põhipalga statistiline võrdlus aastatel 2012 ja 2013

Analüüsi eesmärk on leida, milliste töötajagruppide keskmine täiskoormusel põhipalk on aasta jooksul muutunud ning kas töötajagruppides, kus palk on muutunud, on see tõusnud või langenud. Selleks teostati analüüs, hoides katseviisilise veamäära alla olulisuse nivoo 0,05.

Keskmete võrdlemiseks kasutati kahepoolset t-testi. Kasutades t-testi peab olema täidetud järgnev eeldus:

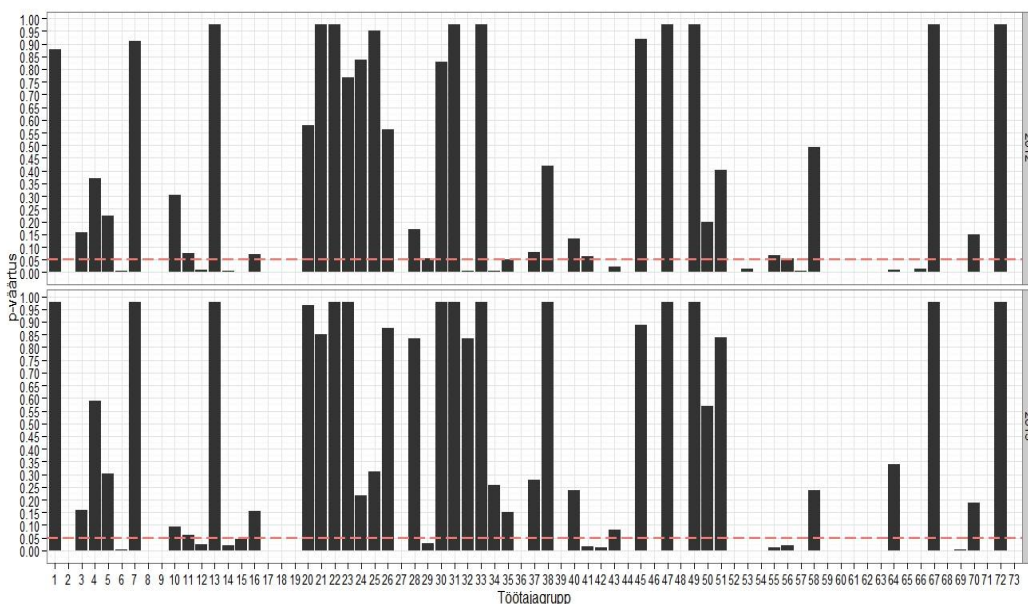
1. Uuritav muutuja on normaaljaotusega mõlemas võrreldavas grupis.

Eelduse kontroll teostati statistikaprogrammiga R, vastav kood on esitatud lisas 3 (lk 42). Normaaljaotuse eelduse kontrollimiseks teostati iga töötajagrupi jaoks ning aastate 2012 ja 2013 puhul eraldi Shapiro-Wilki testi kontrollimaks, kas uuritav tunnus on normaaljaotusega. Shapiro-Wilki testi korral on nullhüpoteesiks uuritava tunnuse normaaljaotus ning alternatiivseks hüpoteesiks see, et uuritav tunnus ei ole normaaljaotusest. Normaaljaotuse kontrollimiseks teostati kokku 146 Shapiro-Wilki testi. Joonisel 4.5 (vt järgmiselt leheküljelt) on esitatud testimisel saadud p -väärtused. Joonisel on punase joonega märgitud meid huvitav olulisuse nivoo (0,05). Gruppide korral, mille p -väärtused on suuremad kui 0,05, jäädakse normaaljaotuse eelduse juurde, teiste gruppide puhul mitte. T-testide sooritamiseks on vajalik, et normaaljaotuse eeldus oleks täidetud mõlemal aastal. Jooniselt on näha, et eeldus on täidetud 12 grupi puhul.



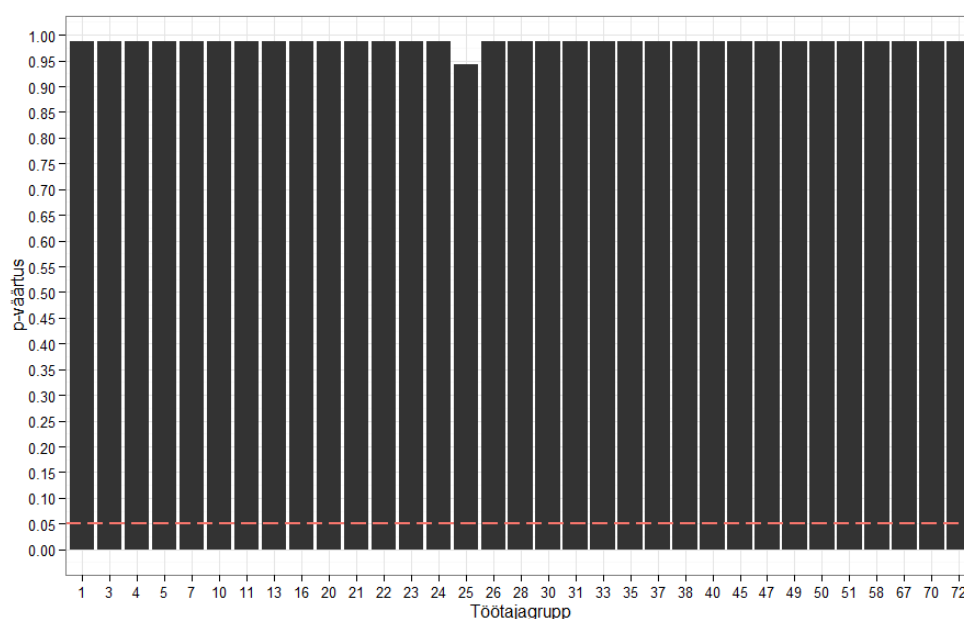
Joonis 4.5. Normaaljaotuse kontrollimisel saadud p -väärtused (kasutamata mitmese testimise meetodeid)

Järgnevalt peame aga kasutama meie eelduse kontrollimisel mitmese testimise meetodit, et tulemus oleks usaldusväärne. Selleks rakendame eelnevalt leitud p -väärtustele Hommeli mitmese testimise meetodit. Hommeli kohandatud p -väärtused on esitatud joonisel 4.6. Jooniselt näeme, et tänu Hommeli meetodi kasutamisele hoiame ära 21 valepositiivse tulemuse vastuvõtmise ning saame 33 töötajagrupi korral jääda nullhüpoteesi juurde.



Joonis 4.6. Normaaljaotuse kontrollimisel saadud Hommeli kohandatud p -väärtused

Lisaks kontrollime ka gruppide dispersioonide võrdsust, milleks teostame võrdsete dispersioonide F-testi. F-test teostatakse iga keskmiste võrdlemise paarile. See tähendab, et teostatakse test võrdlemaks töötajagruppide dispersioone aastatel 2012 ja 2013. F-testi nullhüpotees väidab, et dispersioonid on võrdsed ning alternatiivne hüpotees, et dispersioonid on erinevad. Kontrollimaks dispersioonide võrdsust, teostati 33 F-testi ja saadud p -väärtustele rakendati Hommeli mitmese testimise meetodit. Joonisel 4.7 on esitatud Hommeli meetodil saadud kohandatud p -väärtused.



Joonis 4.7. Dispersioonide võrdlemisel saadud kohandatud p -väärtused

Jooniselt näeme, et kõikide gruppide puhul jääme nullhüpoteesi juurde, mis tähendab, et kõikides töötajagruppide paarides on dispersioonid võrdsed.

Keskmete võrdlemiseks on vajalik iga töötajagrupi jaoks, kus on täidetud normaaljaotuse eeldus, testida hüpoteesi H_i . Püstitatud hüpoteesid on järgmised:

$$H_{0i}: \mu_{1i} = \mu_{2i},$$

$$H_{1i}: \mu_{1i} \neq \mu_{2i},$$

kus i on indeksite hulk, mis vastavad töötajagrupi ID-le, kus normaaljaotuse eeldus on täidetud, μ_{1i} on i -nda töötajagrupi 2012. aasta keskmine täiskoormusel põhipalk ning μ_{2i} on i -nda töötajagrupi 2013. aasta keskmine täiskoormusel põhipalk.

Analüüs teostati statistikaprogrammiga R, vastav kood on esitatud lisas 3 (lk 42). Analüüsi tulemusena koostati tabel 4.1, kus esitatakse testide p -väärtused kasutamata mitmese testimise meetodeid ning samuti ka kohandatud p -väärtused kõiki eelnevalt kirjeldatud mitmese testimise meetodeid kasutades. Tabelis 4.1 on toodud p -väärtused, mis on saadud kasutades kahepoolset t -testi, arvestades, et dispersioonid on võrdsed, võrdlemaks keskmisi põhipalkasid aastatel 2012 ja 2013 ning seda iga töötajagrupi jaoks eraldi.

Esmalt analüüsitakse tulemusi, mis on saadud mitmese testimise meetodeid kasutamata. Tabelist võib näha, et grupi 70 puhul peaks vastu võtma alternatiivse hüpoteesi ehk väitma, et keskmine põhipalk on aastaga muutunud. Grupile 70 vastav p -väärtus on 0,028. Teiste töötajagruppide puhul on p -väärtused suuremad kui 0,05 ning seega jääme nende testide puhul nullhüpoteesi juurde ehk selle juurde, et keskmised aastate lõikes ei erine.

Tabel 4.1. Tartu Ülikooli töötajagruppide keskmiste põhipalkade võrdlusel saadud p -väärtused ja kohandatud p -väärtused

Töötajagrupp	Esialgne p -väärtus	Bonferroni	Dunn-Šidak	Bonferroni-Holm	Simes-Hochberg	Hommel
1	0,429	1,000	1,000	1,000	0,995	0,995
3	0,985	1,000	1,000	1,000	0,995	0,995
4	0,844	1,000	1,000	1,000	0,995	0,995
5	0,723	1,000	1,000	1,000	0,995	0,995
7	0,957	1,000	1,000	1,000	0,995	0,995
10	0,855	1,000	1,000	1,000	0,995	0,995
11	0,962	1,000	1,000	1,000	0,995	0,995
13	0,701	1,000	1,000	1,000	0,995	0,995
16	0,963	1,000	1,000	1,000	0,995	0,995
20	0,516	1,000	1,000	1,000	0,995	0,995
21	0,851	1,000	1,000	1,000	0,995	0,995
22	0,995	1,000	1,000	1,000	0,995	0,995
23	0,811	1,000	1,000	1,000	0,995	0,995
24	0,888	1,000	1,000	1,000	0,995	0,995
25	0,184	1,000	0,999	1,000	0,995	0,995
26	0,401	1,000	1,000	1,000	0,995	0,995
28	0,275	1,000	1,000	1,000	0,995	0,995
30	0,210	1,000	1,000	1,000	0,995	0,995
31	0,657	1,000	1,000	1,000	0,995	0,995
33	0,705	1,000	1,000	1,000	0,995	0,995
35	0,764	1,000	1,000	1,000	0,995	0,995
37	0,921	1,000	1,000	1,000	0,995	0,995
38	0,655	1,000	1,000	1,000	0,995	0,995
40	0,225	1,000	1,000	1,000	0,995	0,995
45	0,729	1,000	1,000	1,000	0,995	0,995
47	0,940	1,000	1,000	1,000	0,995	0,995
49	0,518	1,000	1,000	1,000	0,995	0,995
50	0,626	1,000	1,000	1,000	0,995	0,995
51	0,559	1,000	1,000	1,000	0,995	0,995
58	0,870	1,000	1,000	1,000	0,995	0,995
67	0,991	1,000	1,000	1,000	0,995	0,995
70	0,028	0,932	0,611	0,932	0,932	0,932
72	0,196	1,000	0,999	1,000	0,995	0,995

Analüüsides mitmese testimise meetoditega saadud tulemusi, näeme, et kõikide meetodite puhul jääme kõikide töötajagruppide korral nullhüpoteesi juurde. Seega ei saa tõestada, et töötajagruppides oleks keskmine täiskoormusel põhipalk aastaga muutunud.

Mitmese testimise meetodeid kasutades ei võetud alternatiivset hüpoteesi vastu ühegi töötajagrupi jaoks, mitmese testimise meetodeid kasutamata kinnitati statistiliselt oluliseks erinevus grupi 70 põhipalkades. Seega mitmese testimise meetodid aitasid ära hoida ühe valepositiivse tulemuseni jõudmise.

Võrreldes mitmese testimise meetodeid selle analüüsi tulemuste põhjal, näeme, et tulemused väga palju ei erine – tulemus on kõikide meetodite korral sama, erinevad ainult p -väärtused. Bonferroni meetod, kui kõige konservatiivsem meetod, annab peaaegu kõigi testide korral p -väärtusteks ühed, ainult grupi 70 korral on p -väärtus võrdne 0,932. Teame, et testi võimsused Bonferroni meetodi ja Bonferroni-Holmi meetodi korral erinevad märgatavalt juhul, kui mingi osa p -väärtustest on 0-lähedased. Selle analüüsi korral meil sellist olukorda pole ning Bonferroni ja Bonferroni-Holmi meetodid annavad sama tulemuse. Dunni-Šidaki meetodi korral on ainukesed erinevuskohad gruppide 25 ja 72 tulemused, kus vastavad p -väärtused on 0,999 ning 0,999. Hommeli ja Simesi-Hochbergi meetodite kui kõige võimsamate meetodite korral on grupi 70 korral p -väärtus võrdne 0,932, kõikide teiste testide korral p -väärtused võrdsed 0,995. Simesi-Hochbergi ja Hommeli meetodite võimsuse erinevus selle analüüsi korral välja ei tule.

Kasutatud kirjandus

1. Dunn, O. J. 1959. Confidence intervals for the means of dependent, normally distributed variables. *Journal of the American Statistical Association*, Vol. 54, No. 287 (Sep., 1959), 613–621.
2. Dunn, O. J. 1961. Multiple Comparisons Among Means. *Journal of the American Statistical Association*, Vol. 56, No. 293 (Mar., 1961), 52–64.
3. Hochberg, Y. 1988. A Sharper Bonferroni Procedure for Multiple Tests of Significance. *Biometrika*, Vol. 75, No. 4 (1988), 800–802.
4. Holm, S. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, Vol. 6, No. 2 (1979), 65–70.
5. Hommel, G. 1986. Multiple Test Procedures for Arbitrary Dependence Structures. *Metrika*, Vol. 33, 1986, 321–336.
6. Hommel, G. 1988. A Stagewise Rejective Multiple Test Procedure Based on a Modified Bonferroni Test. *Biometrika*, Vol. 75, No. 2, 383–386.
7. Šidak, Z. 1967. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association*, Vol. 62, No. 318 (Jun., 1967), 626–633.
8. Wright, S. Paul. 1992. Adjuste P-Values for Simultaneous Inference. *Biometrics*, Vol. 48 (Dec., 1992), 1005-1013
9. The Economist, 2013. Trouble at the Lab. *The Economist*, 9. Oct, 21–24.

Lisa 1. R-i kood näite jaoks

```
#R-i kood näite jaoks
set.seed(123)
#genereeritakse 20000 juhuslikku suurust standardiseeritud normaaljaotusest
x=rnorm(20000)

#jagatakse juhuslikud suurused 200-ks valimiks, millel igal on 100 vaatlust
maatriks=matrix(x, nrow=200, ncol=100)

#moodustatakse andmetabel 200-st valimist nimedega X1, X2, X3, ...
jaotatud=as.data.frame(split(maatriks,rep(1:200, each=100)))

#teostatakse t-testid valimipaaridele X1&X2, X3&X4, X5&X6, ...
testid=lapply(seq(1,length(jaotatud),by=2),function(x){t.test(jaotatud[,x],
jaotatud[,x+1])})

#väljastatakse ainult p-väärtused
pvals = sapply(testid, function(x){x$p.value})

#Joonis
qplot(pvals,      xlab="p-väärtused",      ylab="",      geom="bar")+theme_bw()+
geom_vline(xintercept      =      0.05,      colour="#F8766D",      size=1,
linetype="longdash")+scale_x_continuous(breaks=seq(0, 1, 0.05))
```


Lisa 2. SAS-i kood näite jaoks

```
/*Genereerin 20000 juhuslikku suurust standadiseeritud normaaljaotusest */
data andmed;
call streaminit(123);
do i =1 to 20000;
esimene=rand('normal');
output;
end;
/*Loon tunnuse, mille abil jaotada genereeritud suurused 100 võrdlemisgrupiks */
data andmed2;
array a{20000} (100*1:200);
output;
run;
proc transpose data=andmed2
                out=andmed2_transposed;
run;
proc sort data=andmed2_transposed;
by COL1;
run;
/*Loon tunnuse, mille abil jaotada genereeritud suurused 200 valimiks */
data andmed3;
array a{20000} (200*1:100);
output;
run;
proc transpose data=andmed3
                out=andmed3_transposed;
run;
proc sort data=andmed3_transposed;
by COL1;
run;
data andmed3;
set andmed3_transposed;
rename COL1=grupp;
run;
/*Ühendan genereeritud juhuslikke suurustega andmed, valimiteks määrava tunnuse ning
võrdlemisgruppideks määrava tunnuse*/
data uus;
merge andmed andmed2_transposed andmed3;
run;
/*Teostan 200 valimiga 100 t-testi võrreldes igas 2 valimikeskmist ning väljastan saadud p-
väärtused */
proc ttest data=uus;
by grupp;
class COL1;
var esimene;
ods output Ttests=ttest_output;
run;
/*Jätan alles p-väärtused, kus eeldatakse dispersioonide võrdsust*/
data uus2;
set Ttest_output;
if method="Pooled" ;
run;
data loplik;
set uus2;
rename Probt=raw_p;
run;
/*Arvutan Hommeli kohandatud p-väärtused*/
proc multtest INPVALUES=loplik hommel ods=results;
run;
```

Lisa 3. R-i kood analüüsi jaoks

```
#R-i kood analüüsi jaoks
#Eelduste kontroll
#Normaaljaotuse kontrollimine

#Shapiro-Wilkinsoni test
normal_p_vaartus=ddply(andmed2,c("Grupi_ID","Aasta"),function(x){data.frame(p_vaartus=shapiro.test(x$sissetulek)$p.value)})

#Hommeli meetod
adjust_normal=data.frame(Aasta=normal_p_vaartus$Aasta,Grupp=normal_p_vaartus$Grupi_ID,adjust=p.adjust(normal_p_vaartus$p_vaartus, "hommel"))
#Jätan alles vaid eeldusele vastavad töötajagrupi
vajalik=andmed2[andmed2$Grupi_ID %in%
c(1,3,4,5,7,10,11,13,16,20,21,22,23,24,25,26,28,30,31,33,35,37,38,40,45,47,49,50,51,58,67,70,72), ]

#Joonis
qplot(data=adjust_normal,x=Grupp,xlab="Töötajagrupp",ylab="p-väärtus",
y=adjust,facets=Aasta~.,geom="bar")+geom_hline(yintercept=0.05,colour="#F8766D",
size=1, linetype="longdash")+scale_y_continuous(breaks=seq(0, 1, 0.05))+theme_bw()

#Dispersioonide võrdlemine
#Dispersioonide võrdlemise F-test
variance_p_vaartus=ddply(vajalik,"Grupi_ID",
function(x){data.frame(p_vaartus=var.test(x[x$Aasta=="2012"],$sissetulek,
x[x$Aasta=="2013"],$sissetulek)$p.value)})

#Hommeli meetod
adjust_variance=data.frame(Grupp=variance_p_vaartus$Grupi_ID,adjust=p.adjust(variance_p_vaartus$p_vaartus, "hommel"))

#Joonis
qplot(data=adjust_variance,x=Grupp,xlab="Töötajagrupp",ylab="p-väärtus", y=adjust,
geom="bar")+geom_hline(yintercept=0.05,colour="#F8766D",size=1,linetype="longdash")
+scale_y_continuous(breaks=seq(0,1,0.05))+theme_bw()

#ANALÜÜS
#Kahepoolne t-test võrdsete dispersioonidega
analyys=ddply(vajalik,"Grupi_ID",function(x){data.frame(pvaartus=t.test(x$sissetulek~x$Aasta, var.equal=T)$p.value)})

#Bonferroni meetod
vaartused1=p.adjust(analyys$pvaartus, "bonferroni")
#Bonferroni-Holmi meetod
vaartused2=p.adjust(analyys$pvaartus, "holm")
#Simesi-Hochbergi meetod
vaartused3=p.adjust(analyys$pvaartus, "hochberg")
#Hommeli meetod
vaartused4=p.adjust(analyys$pvaartus, "hommel")
#Dunni-Šidaki meetod
sidak = function(p){
tulemus = 1-(1 - p)**(length(p))
return(tulemus)
}
vaartused5=sidak(analyys$pvaartus)
```

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, _____ Liina Tamme _____.

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
_____ Mitmese testimise probleem _____.

mille juhendaja on _____ Marju Valge _____,

- 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. Olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, **05.05.2014**